Gender & Breeding Initiative

CGIAR

Working Paper 1

Gender and social targeting in plant breeding

Alastair Orr, Cindy M. Cox, Yating Ru, Jacqueline Ashby

2 0 1 8

CGIAR Gender and Breeding Initiative Working Paper

Editing: Gary Harrison

Design: CPAD-CIP, Lima, Peru

Layout: Zandra Vasquez, RTB-CIP, Lima, Peru

Contact:
CGIAR Gender and Breeding Initiative
International Potato Center (CIP)
Apartado 1558, Lima 12, Peru
rtb@cgiar.org • www.rtb.cgiar.org

ISSN 2309-6586
DOI+ISBN

Correct citation:
Orr, A., Cox, C.M., Ru, Y. and Ashby, J. 2018. Gender and social targeting in plant breeding. Lima (Peru). CGIAR Gender and Breeding Initiative. GBI Working Paper. No. 1. Available online at: www.rtb.cgiar.org/gender-breeding-initiative

Author affiliation:
Alastair Orr: Honorary Research Fellow, ICRISAT Nairobi, Kenya
Cindy Cox: IFPRI, Washington, USA
Yating Ru: IFPRI, Washington, USA
Jacqueline Ashby: International Development Consulting

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

AEZ	Agro-ecological zones
DHS	Demographic and Health Surveys
ESA	East and southern Africa
FHH	Female-headed household
GIS	Geospatial information systems
IFPRI	International Food Policy Research Institute
IIASA	International Institute for Applied Systems Analysis
ISA	Integrated Surveys on Agriculture
LSM	Life Style Measurement
LSMS	Living Standards Measurement Study
ME	Mega environments
SPAM	Spatial Production Allocation Model
SSA	Sub-Saharan Africa
STP	Segmenting-Targeting-Positioning
TPEs	Target population environments
WEAI	Women's Empowerment in Agriculture Index

PREFACE

For plant and animal breeders to meet users' needs, they need to understand the priorities that women and men assign to genetically determined traits – such as taste, color, size and shape. Many CGIAR breeding programs know that if they overlook traits important to women users, this can aggravate household food insecurity and poverty. However, breeding programs still don't have enough practical methods and tools to help them decide how to be more gender-responsive and consider gender differences in breeding schemes. Tackling this knowledge gap is urgent if CGIAR Centers and Research Programs (CRPs) are to achieve the targets for gender equality defined in the CGIAR Strategy and Results Framework.

In response, the CGIAR Gender and Breeding Initiative was launched in 2017, building on a strategy developed by an interdisciplinary group of breeders and social scientists who came together in 2016 as part of a workshop on "Gender, Breeding and Genomics" convened by the CGIAR Gender Network (which has now evolved into the CGIAR Collaborative Platform for Gender Research led by the CGIAR Research Program on Policies, Institutes and Markets).

The Initiative brought together a broader group of scientists in October 2017 to build on this earlier work and develop recommendations for practical ways to improve the gender-responsiveness of breeding programs; evidence-based methods and tools for gender-responsive targeting, implementation of breeding activities and linkage with variety dissemination; and support a community of practice for active sharing and development of methods and tools.

This working paper is part of a series of knowledge products designed to share the outputs from the 2017 "Innovation in Gender-Responsive Breeding" workshop, and to share the Initiative's collective knowledge more widely across CGIAR and partner breeding programs.

The Initiative is coordinated by the CGIAR Research Program on Roots, Tubers and Bananas and the International Potato Center, with funding support by CGIAR Fund Donors.

Graham Thiele, Director
CGIAR Research Program on Roots, Tubers and Bananas (RTB)

Stefania Grando, Chair
Honorary Fellow, ICRISAT and International Consultant
Chair, Workshop Organizing Committee

ACKNOWLEDGMENTS

We are grateful to Peter Kulakow, Yilma Kebede, Mike Olsen, and Michel Ragot for discussions on plant breeding; to Hale Ann Tufan and Bela Teeken for discussions on cassava; and to Joab Ouma for discussion on precooked beans. Comments from participants at the Workshop on Gender-Responsive Breeding were helpful in revising the paper.

The views expressed in this paper are those of the authors and should not be attributed to any organization with which they are affiliated.

EXECUTIVE SUMMARY

Gender and social targeting can improve the relevance and effectiveness of plant-breeding programs serving resource-poor farmers, traders, processors, and consumers. Generally, these breeding programs have limited information about their clients, which makes it difficult to prioritize breeding objectives. As a result, products from these breeding programs may not meet the needs of their intended users.

We argue that plant breeding for resource-poor farmers, sellers, and processors requires a marketing approach. We show how the Segmenting-Targeting-Positioning (STP) framework from consumer marketing can be adapted for gender and social targeting in these programs. First, Segment the market, or identify groups of consumers with homogeneous preferences ("market segments"). Second, Target those market segments that meet the programs' equity objectives, are big enough to justify the investment, and whose preferences match physical traits. Third, Position new products in the market by showing how these new products meet the preferences of their intended users.

The STP framework is broken down into eight logical steps which provide a checklist for gender and social targeting. The result is a "customer profile" (just like a breeders' "product profile"), which combines demographic, behavioral, and geographic variables with a set of trait preferences to describe a market segment. A customer profile gives the program a clear picture of whom the program is breeding for, the expected number of customers, and why they prefer specific traits.

To prioritize breeding objectives, breeders must have an accurate picture of the relative size and social character of different client groups. Currently, information about these clients and their trait preferences is based on small-scale studies, which makes it difficult to set breeding priorities at the national or regional level. But the growing number and availability of large datasets make it possible to define growers and crop utilization on a much bigger scale. We inventory large datasets, identify a minimum dataset of biophysical and socioeconomic variables, and show how these variables can be layered for gender and social targeting at the national level. Datasets include the Living Standards Measurement Study–Integrated Surveys on Agriculture (LSMS–ISA), the Women's Empowerment in Agriculture Index (WEAI), and the Demographic and Health Surveys (DHS) Program. We use the example of cassava in Nigeria to illustrate how these datasets can help breeding programs incorporate gender into their customer profiles.

Finally, we suggest ways to improve the design of gender and social targeting studies in order to enhance their relevance for plant-breeding programs:

1. Breeders and social scientists should agree on the problem.
2. Use the STP framework to see where a targeting study fits within the bigger picture of developing a targeted breeding program. We suggest how to operationalize the framework as a checklist or as a matrix.
3. Use large datasets that can give the program information about potential target groups at national scale. Information about trait preferences from small-scale studies can be linked to these datasets.
4. Use mixed methods. Quantitative methods are needed to identify market segments and to develop customer profiles, but qualitative methods give insights into the reasons for trait preferences. We provide examples of the methods and tools available for each stage of the STP framework.

Gender and social targeting in plant breeding

1. INTRODUCTION

Public investment in international aid aimed at sustainable development for the poor has recently been channeled into a renewed commitment to agricultural research centered on plant and animal breeding. The expectation is that breeding in the twenty-first century can tap opportunities for novelty comparable to the Green Revolution of the 1960s. In the past, however, the benefits of this multimillion dollar investment were unevenly distributed among different groups of poor people; women in particular, have been overlooked and disadvantaged because social targeting by breeding programs is imperfect (Lipton and Longhurst 1989; Lipton 2007). This is costly and counterproductive. Despite renewed attention to the importance of social and gender equality for development, published strategies seldom explicitly identify gender-differentiated social groups in their stated targets. The need for improvement in social targeting remains largely ignored in the recent renewal of breeding programs intended to benefit the poor.[1] This paper shows how this shortcoming can be overcome.

The proposed approach to gender-responsive, social targeting for use in public sector breeding is based on several premises. First, our concern is with social targeting for breeding programs that aim to benefit small, resource-poor producers in one or more developing countries. Such programs may also breed for large, commercial or industrial-scale producers as a means to reduce food costs for the urban and rural poor; but this strategy is not our focus. The relevance of gender differences for demand for breeding products by large-scale commercial producers and consumers fully integrated into markets can be identified from readily observable market signals. Our focus is on situations in which gender differences in demand are obscured by imperfect markets.

Second, this paper focuses on crop breeding. Although we began work on this paper with the intention of covering both crop and animal breeding, important differences between the two made it evident that social targeting for animal breeding requires a separate discussion. Nonetheless, many features of the approach proposed for social targeting have value for animal breeders.

Third, we make a distinction between functional and transformative approaches to gender equality. A functional approach takes gender differences among small producers into account only when the delivery of relevant breeding products to both men and women users is essential for achieving desired levels of adoption and impact. A transformative approach to gender equality takes gender

[1] For example, the UN Sustainable Development Goal concerned with agriculture only specifies "small-scale food producers, particularly women, indigenous peoples, family farmers, pastoralists and fishers." CGIAR research to reduce poverty and improve nutrition and ecosystem services prioritizes the development of improved high-yielding and stress-resistant crop varieties, livestock, and fish breeds and expansion of the overall benefits accruing from farm systems, in part through the creation of more opportunities for women worldwide (CGIAR 2015). Strategy does not, however, go beyond "women" to specify socially defined target groups with a gender dimension.

differences into account in order to reduce the gender gap between men and women producers. This may require breeding to target certain types of women producers and develop breeding products specifically with a transformative objective. The majority of breeding programs that aim to be gender responsive have functional rather than transformative objectives. We point out key differences in approach to targeting that may arise when a breeding program needs to decide whether to take a functional or a transformative approach.

1.1 Social and Gender Targeting When Breeding for the Poor

Breeding programs for resource-poor farmers have limited information about their clients. In the private sector, information about clients and their trait preferences is provided by market research, and the accuracy of this information is tested by product sales and by market share. Public sector breeding programs intended to benefit resource-poor farmers do not have these advantages. This makes it difficult to prioritize breeding objectives and to develop the right products. As a result, breeding intended to benefit resource-poor farmers may not match the needs of their intended users.

Typically, resource-poor farmers, traders, and consumers operate in imperfect markets for which there is limited knowledge of actual or potential demand for varieties. Public sector programs for breeding improved crops or livestock for the poor customarily cope with this uncertainty by relying on agricultural extension, participatory plant breeding, and household surveys for information about demand. One problem with this information is its scale: often it is derived from small, self-selected groups of informants and poorly designed samples, which makes it difficult to generalize results. Agricultural household surveys, although usually based on formal sampling, are typically designed to collect detailed information about specific crops or farming systems. These studies do not usually set out to characterize crop growers or farm systems' inhabitants, to encompass a wide range of local contexts and cultures, or to provide socioeconomic measures that are internationally comparable and incorporate a gendered dimension. Consequently, breeding programs typically lack representative information with a gender dimension about demand for different breeding products at the scale of a national or regional population of clients for their products. Breeding objectives are set without accurate information about the size and relative importance of different client groups.

Scale is not the only problem. In the absence of reliable, representative information about the characteristics and preferences of different client groups, breeding objectives are set primarily in relation to geography instead of demography. Information about production constraints of a crop in an area are always considered—temperature, rainfall, soil type, for example. These production constraints are important, but understanding socioeconomic constraints and demand remains vital for adoption. Geography is not enough. Breeding objectives need to be set with a combination of geographic and social targeting that together provides a profile of a given client population, their trait preferences, and end-uses of the crop or animal breed in question, within a geographic production domain. Demographic characteristics such as differences in resource ownership, gender, age, and ethnicity need to be considered when breeding objectives are set, as these are highly correlated with social drivers of trait preferences, such as ability to purchase farm inputs or to market surpluses. In sum, scale and social targeting both matter. To prioritize breeding objectives and develop products that will be adopted, breeders must have an accurate picture of the relative size and social character of different client groups, in addition to their production constraints. Breeders accept that prioritization means trade-offs among different traits. What needs to be added is that prioritization also involves making explicit trade-offs among different beneficiary groups when these have contrasting or even competing trait preferences.

Making breeding programs more customer driven requires a marketing approach (Sumberg and Reece 2004; Sumberg et al. 2013). In the world of business marketing, questions about how to identify different types of customer, their distinctive preferences, and matching products to fit their needs are all familiar problems and the subject of intensive research. In the world of plant breeding, private seed companies use a marketing approach. They operate in a competitive market where the success of their products depends on their knowledge of customer needs, market demand, and profitability. Their competitive edge relies on good market research. The final decision to release a new crop variety is made not by plant breeders but by the marketing team. Both business marketing and the experience of private seed companies can offer useful guidance to a plant breeding program intended to benefit the poor.

Central to a marketing approach is the concept of targeting. “Target groups” are a familiar but ill-defined concept in agricultural research. In marketing, however, targeting means the selection of a *homogeneous group of consumers with a unique set of preferences.* In marketing, targeting is just one stage in a wider process of developing and selling a new product. First, the company must define the market in which it operates and segment that market into homogeneous consumer groups with distinctive preferences. Only then is the decision made to target a specific group. In this paper we propose that in the same way, breeding programs must first have a clear picture of the demand for a particular crop and its different end-uses, as well as different types of users, before it can target the poor and subgroups within the population of poor people, by understanding their unique preferences.

Without gender and social targeting, there is a danger that breeding programs may end up serving the wrong clients. For example, IITA’s breeding program for cassava in Nigeria has emphasized yield and dry matter content. Demand for starch from industrial users, however, is still limited and the main market for cassava is for *gari*, a fermented flour. This may have contributed to a low adoption of improved varieties that have high starch content but lack the quality traits preferred for *gari*. Furthermore, cassava for starch is grown mostly on medium-sized farms, whereas cassava for *gari* is grown on smaller farms and usually processed by women. This highlights the importance of gender and social targeting in setting breeding objectives. Better information about growers, processors, and end-uses could have led to a different set of breeding objectives.

Targeting thus depends on information—the size of the market, the number of customers, and their preferences. Consumer marketing uses large databases for this purpose. For example, the Target Group Index has information on consumers worldwide, including 10 African countries. Breeding programs intended to benefit poor end-users of a crop or animal breed in multiple developing countries do not have comparable datasets. Nonetheless, there are now large datasets covering multiple countries or regions that can help fill this information gap. These include multicountry household expenditure surveys; integrated household surveys that capture agriculture, nutrition, and expenditure, production, and welfare variables; and demographic and health surveys. Often these surveys are repeated at regular intervals, and some are panel surveys that revisit the same households periodically. In addition, datasets are available for some of the same countries or regions that provide biophysical information, including rainfall and temperature, length of growing period, and soils. In combination, these datasets provide a rich source of information for social targeting.

Information from large-scale datasets can be used for targeting in several ways. First, it can reveal the market scale and composition (differentiating homogeneous client groups) for breeding products (i.e., the number of end-users of different types and their location). Second, when linked with information on clients' trait preferences obtained from microstudies, this information will allow the results of these studies to be generalized to the wider population. This can give programs a better picture of their clients' needs, which in turn helps programs to prioritize products with the highest potential demand. So far, however, the exploitation of large-scale, multicountry datasets for characterizing end-users for breeding programs has been limited.

1.2 Objectives of the Paper

This paper addresses the following three overarching research questions:

1. What do breeding programs need to know about the demographic, sociocultural, and market conditions of their intended target population(s) in order to target them effectively, taking into account the intersection of gender with other sociocultural, economic, and geographic characteristics of importance?
2. How can this analysis be conducted either with available data or data that can be rapidly generated? How can breeding programs obtain reliable and representative information about gender-differentiated trait preferences that is correlated with the important demographic, sociocultural, and market information?
3. What should a study to improve gender and social targeting in a breeding program look like? What conceptual framework, research questions, sources of data, and tools might such a study use?

This paper is written for plant and animal breeders and social scientists supporting breeders who want to know how social and gender targeting can improve priority-setting within their program. It sets out a framework and identifies datasets that social scientists can use to provide breeders with the information they require for gender and social targeting. The emphasis is not on theory but on practicalities—research questions, information, and tools. The aim is to provide plant and animal breeders with enough information for them to start a dialogue with social scientists about how to apply social and gender targeting within their breeding programs.

The paper comprises five sections. Following this introduction, section 2 presents a conceptual framework for gender and social targeting. Section 3 provides an inventory of useful datasets, a list of key targeting variables, and a case study applied to cassava in Nigeria. Section 4 identifies general principles for the design of gender- and social-targeting studies. The final section summarizes the key messages of the paper.

2. WHAT DO BREEDING PROGRAMS NEED TO KNOW?

In order to target beneficiary groups among the poor, a breeding program needs to use a framework for targeting with a sequence of stages and steps compatible with the breeding cycle. This paper follows the Segmentation-Targeting-Positioning (STP) framework used in consumer marketing (Figure 1). In target marketing, the seller distinguishes the major market segments, targets one or more of these segments, and develops products tailored to each selected segment (Kotler 2000). Despite its critics, it remains the textbook approach to consumer marketing.

Figure 1. The STP framework.

Stage	Description	Data Required
	"S": Segmentation	
1. Define the market	Generic market: aggregate market for a product Relevant market: boundary to guide breeding program Defined market: existing customers, potential customers	Target countries Agro-ecosystems Area planted to crop Value chains for crop End-uses for crop
2. Select bases for segmentation	Geographic (where?) Demographic (who?) Behavioral (why?)	Region, state Age, marital status, gender, ethnicity, income, occupation, consumption/sale End-uses, trait preferences
3. Validate customer segments	Measurable Substantial Accessible Differentiable (conceptually distinct and respond differently to market stimuli) Actionable (program can be designed to serve the market) Stable (sufficiently stable to justify investment in market)	Size Purchasing power Profitability Growth rate Location Distance to market
4. Construct customer profiles	Develop a socioeconomic profile of customer segments:	Additional socioeconomic variables not used in Steps 2 and 3
	"T": Targeting	
5. Evaluate market attractiveness of segments	Exclude segments with low numbers of poor producers	Number of poor, non-poor producers
6. Identify which and how many segments should be targeted	Compare segments with resource-poor producers	Size of segments, growth rate, number of resource-poor producers, location, distance to market
RESULT	Top priority target segments selected. Customer profile for each segment selected to be targeted—a *homogeneous group of consumers with a unique set of preferences*	
	"P": Positioning (Market)	
7. Develop positioning strategy	Differentiate the product to give it a competitive advantage in the target market Assign a hierarchy of attributes that consumers use in selecting a brand	
8. Design appropriate marketing mix to communicate positioning	The 4 Ps: product, price, promotion, place	
RESULT	Product profile keyed to one or several target customer profiles	

Sources: Kotler (2000); Simkin and Dibb (1998); Weinstein (2006)

The STP framework sets out a logical three-stage sequence for breeding programs:

1. Segment, or characterize the population into homogeneous groups (Steps 1–3)
2. Target, or identify the specific customer groups and their preferences (Steps 4–6)
3. Position the new product in the market by tailoring communication with these users (Steps 7 and 8).

Each of the three stages in the framework can be divided into steps, with eight steps in total. This paper focuses on Steps 1–3 (*segmentation*) and Steps 4–6 (*targeting*). Steps 7 and 8 (*market positioning*) are important for certain products with new or unusual attributes. Marketing expertise is required to identify the most effective way to win consumer acceptance for these products; however, the main focus of this paper is on segmentation and targeting.

2.1 SEGMENTING THE MARKET

Market segmentation involves viewing a heterogeneous market as a number of smaller homogeneous markets, in response to differing preferences, attributable to the desires of consumers for more precise satisfaction of their varying wants. (Smith 1956)

Breeding programs make products for uptake by farmers, fishers, or livestock keepers. Hence, the customer for a breeding program defined in this way is not the consumer but the producer. Producers are not a homogeneous group; they can be split into different "market segments," or subgroups of producers with similar preferences. As defined by the GBI (2017) glossary, market segment is "a group of producers having a relatively homogeneous demand for a commodity (here crop varieties or animal breeds)."

When using the STP framework in the context of a breeding program, it is important to distinguish between customers, who are the target, and beneficiaries. The market segments targeted as customers for the purposes of breeding are the producers. Producers are the customers for the program because they make the decisions on whether to adopt the product.

The "beneficiaries" of a breeding program are a different set from customers. For example, consumers are not termed customers. Instead, they are potential beneficiaries (see definitions in Table 1). Beneficiaries can include service providers (who benefit from increased demand for inputs or credit); processors (who benefit from more consistent supply, higher demand, or by having products that meet their quality standards); and consumers (who benefit from improved nutrition, higher quality, or lower commodity prices). To varying degrees, beneficiaries' preferences influence what producers decide to adopt.

Of course, producers—in particular, the early adopters or producers for home consumption—capture benefits from using a new variety. In the long run, however, for non-export commodities not all producers will benefit if supply of a commodity is increased without a compensating increase in demand, because increased production reduces prices. For this discussion, therefore, the important question for deciding on whom to target is "who will directly adopt the breeding product(s)?"

Table 1. Definitions of terms

Market segment	A group of producers with homogeneous trait preferences for a breeding product, taking into account the gender differentiation of preferences.
Target segment	A market segment selected to be the customers of a breeding program. A target segment may be male, female, or mixed male and female depending on the importance of gender differences for segmentation.
Customer profile	A set of demographic, behavioral, and geographic attributes with a gender dimension, associated with a target segment.
Product profile	A set of targeted attributes that a new crop variety or animal breed must meet to meet the demands of a customer.
Target beneficiary	Anyone who is selected to derive a material benefit from a breeding program according to that program's goals.

Source: Authors

Step 1: Define the market

STP distinguishes between different levels of the market. The first level is the "relevant" market, or the market that is appropriate for the organization given its resources, objectives, and environment (Weinstein 2006). Here, relevance means the geographic scope of the market, the products the organization makes, and the "generic" market to which its products and services are sold. To the extent possible, the relevant market needs to be defined, taking gender into account at each step in segmentation.

This is because the relevant market may be fundamentally different for men than for women, and this needs to be assessed iteratively. It may be that the relevant market segment for achieving a transformative outcome can be defined a priori as a population of women producers with a specific set of constraints that can be solved cost-effectively by breeding. In contrast, the relevant market may encompass both men and women producers facing similar constraints as far as an intervention based on breeding is concerned.

Once the relevant market has been identified, the second level is the "defined" market. This includes both existing customers and new, potential customers that have not yet been reached (the "untapped" market). The defined market is especially interesting from the point of view of inclusiveness and responsiveness to gender because certain types of women producers or other excluded groups may constitute an untapped market.

In the public sector, policy may identify the relevant market for breeding programs concerned with welfare objectives in terms of a broad geography (an agro-ecosystem), a generic set of beneficiaries ("the poor"), and specific crops or animals. In this context, the term market includes the share of the crop that is sold and the share not sold but used for home consumption; however, in this context the defined market is not predetermined. Programs can make choices about which subgroups of the poor they will target (e.g., a program may identify a certain type or class of women producer as a subgroup they will target as a defined market).

Step 2: Select bases for segmenting the market

A basis for segmentation is defined as "a set of variables or characteristics used to assign potential customers to homogeneous groups" (Wedel and Kamakura 1998). In marketing language, therefore, the *bases* for segmentation refer not to single variables but to the composite categories or axes that can be used to define different segments of the market).

Marketing literature identifies three bases relevant for breeding programs (Table 2). The *geographic* base includes the biophysical variables that determine the traits required for adaptation to a specific agro-ecosystem. The *demographic* base includes the gender and social variables that describe the clients of the breeding program. Segmentation will benefit from analysis of the intersection of gender with other social attributes—in particular, wealth, income, or food insecurity—as well as age, ethnicity, or culture. Especially if gender has been identified as a key factor for market relevance, the way gender interacts with other variables relevant for assigning customers to homogeneous groups needs to be understood. This allows segmentation to progress beyond the overly simplified category "women," to a typology differentiating women producers, some of whom (but not necessarily all) have a distinctive demand for breeding products. The *behavioral* base includes the variables that influence consumer wants, such as the benefits they want from the product. "Benefits" in marketing terminology refers to the perceived value or advantage consumers perceive that they receive from a product. Products are usually designed for customers seeking a specific combination of benefits (Haley 1968, 1984). In a breeding program, benefits correspond to trait preferences or what attributes clients (producers, traders, processors) want in a new product.

Table 2. Bases of segmentation and variables for market segments

Base	Variables relevant for a breeding program
Geographic	Agro-ecosystem
	Region
	Principal crops and/or animals Production constraints (e.g., prevalent pests and diseases)
	Distance from market
Demographic, disaggregated by sex of individuals or households	Access to productive resources
	Age
	Farm size
	Use of agricultural laborers vs. unpaid, family farm workers
	Ratios of women to men in (1) household and (2) the area's resident and/or emigrant population
	Ethnicity
	Income and expenditure
Behavioral disaggregated by sex of individuals	Trait preferences
	Technology choice (adopter/non-adopter)
	Awareness or knowledge
	Sale vs. home consumption of product
	Distance to market

Source: Adapted from Kotler (2000).

Step 3: Validate the market segments

This step evaluates the robustness of the market segments, and how well research measures what it claims to measure.

Which bases should we use for segmenting markets? The general rule of criterion validity is that the available independent segmentation variables are closely associated with the dependent criterion of interest, which is usually some aspect of behavior (Tonks 2009). The association may or may not be causal, but the essential requirement is that the independent descriptor variable discriminates the dependent criterion in a useful way. The general rule is that homogeneity is required within segments.

What are the “independent descriptors” required to identify market segments for a plant breeding program? All plant breeding programs use descriptors based on geography or biophysical descriptors. For example, sorghum breeding in the USA is based on adaptation zones based on climate (temperate/ subtropical) and then further subdivided according to rainfall (wetter/drier) and length of growing period. However, biophysical descriptors alone are not sufficient to identify a market segment of resource-poor farmers. First, poor people do not always live in poor environments. Some agro-ecosystems (“marginal environments”) may contain a high share of resource-poor farmers; but the greatest numbers of poor people are usually found in high-potential environments. Second, within any given agro-ecosystem, resources are unequally distributed. Certain social groups may be disadvantaged. Resource-poor farmers (particularly women) may have less fertile land, for example, or poorer access to water. They may also be socially excluded from the more profitable value chains, and lack equal access to extension services or fertilizer. Thus, a breeding program for resource-poor farmers cannot rely primarily on geographic indicators to define this market segment. They are necessary but not sufficient. The required combination of geographic and social indicators is explored further in Section 3 below.

Gender-responsive breeding should include gender as a variable in market segmentation. The market segment of poor producers can be further subdivided into market segments of poor women and poor men. This is important where gender norms mean there are specific so-called women’s crops or women’s animals. However, even if they produce the same crop, it may be relevant to distinguish between women and men, because (1) the desired benefits (traits) may differ or the same traits may be valued quite differently, and (2) women may be disadvantaged in access to resources, material inputs, and knowledge. For example, in Mali sorghum is grown by men and women alike, but women grow sorghum on less fertile, phosphorous-deficient land. Consequently, tolerance to low phosphorus is a requisite trait for female sorghum growers in Mali.

Targeting
Once the market segments have been identified, the program has to decide how many and which segments to target (discussed in Step 6).

Step 4: Construct segment profiles

In this step, each segment is profiled in terms of their socioeconomic characteristics. This is similar to cluster analysis, where clusters are first identified using a set of variables selected for their explanatory power. The clusters are then profiled using a different set of variables used to describe the clusters and explain differences among them. In the same way, the profile of a market segment is not built with the same variables used to define the segment but instead uses additional variables to provide a more comprehensive description. Testing the statistical significance of these additional variables can give further insights into differences between the market segments.

The objective of Step 4 is to construct a customer profile. The customer profile is analogous to the breeders' plant "ideotype" (the ideal product or trait combination) in that it provides a "sociotype" (an ideal customer). The customer profile provides the breeding program with a portrait of a target group in terms of (1) socioeconomic indicators, such as assets, income, and production constraints; (2) specific trait preferences and the reasons for these preferences; and (3) the specific roles they play in the production, sale, and processing of the commodity relevant for the breeding program. A further step is to personalize the segment profile by constructing a *persona*. Personas are fictitious characters created by a retailer to mimic a real customer. They can be useful as a communication tool and help make the profile less abstract. They are a way to visualize the customer profile (see Box 1).

Box 1. A customer profile

A customer profile combines demographic, behavioral, and geographic variables with a set of trait preferences to describe a market segment that will be targeted by the breeding program. We illustrate this process with an example from Mali, where the breeding program for sorghum identified a market segment of female producers that had previously been overlooked as potential customers for improved varieties.

Geography: Sorghum is grown as a staple crop in the Sahelo-Sudanian zone, where annual rainfall is above 600 mm but variability is high, between 600 and 1,200 mm.

Producers: Sorghum is produced by large extended households (10–100+ members) and mainly grown in family fields under the control of the male household head (*chef de famille*) who is in charge of family labor. In addition, women produce sorghum on the individual fields allocated to them by the head. Sorghum on these fields is used to supplement the dietary needs of younger children and to earn cash income.

Trait preferences: Sorghum producers required improved varieties that were sensitive to photoperiod and could adapt to variable rainfall. The latter was particularly important for female producers because they did not own oxen and family fields are ploughed first, so their sorghum plots were often planted late. In addition, producers required improved varieties that were adapted to low soil fertility. Again, this was particularly important for women because they were allocated fields at the end of the rotation where yields were too low to be used as family fields, and they had no access to farmyard manure, which was reserved for family fields. Low phosphorous on women's plots led to late-flowering sorghum and low productivity. Finally, women producers required varieties that were tall and matured early, making them suitable for intercropping with groundnuts. Women producers usually intercropped sorghum with groundnuts because they were responsible for preparing the groundnut sauce that accompanies the main meal.

Gender: Sorghum in Mali provides a very clear example of how social and gender inequality translates into trait preferences. In this example, women's trait preferences are the result of biased access to agricultural inputs (e.g., land, oxen) and biased access to technology (e.g., varieties). Gendered preferences for traits associated with adaptation to low phosphorous and early maturation reflect structural inequality in access to resources, not just differences in gender roles within the household.

Source: https://goo.gl/jDT8wp

Market segments are best constructed using large datasets that are representative at the national level. So far, the exploitation of these datasets for breeding programs has been limited; however, one example illustrates their potential. CIAT has used data from the Life Style Measurement (LSM) consumer database to understand the market for precooked beans in Kenya. First, market demand was broken down into rural and urban, which identified a growing market among urban, middle-class consumers. Second, market demand was subdivided into different consumer segments, based on the level of household income. This identified three consumer segments: the poor (51%) who want affordable, quick-cooking beans, middle-income consumers (44%) who want nutritious beans that are easy to cook and can be cooked with other products like tomatoes, and high-income consumers

(5%) who want trendy bean snacks. Market researchers were “immersed” with these consumer segments to learn their preferences first-hand. CIAT was then able to screen bean varieties and select those that met the needs of each consumer segment (Aseete et al. 2016; Chege et al. 2016; Ouma 2016).

Step 5: Evaluate market attractiveness of segments

Once market segments have been identified, the program has to evaluate which segment(s) are most attractive as a business proposition. The marketing literature identifies five criteria for market attractiveness (Kotler 2000; Littler 1995; Simkin and Dibb 1998).

1. *Measurable*: Can the size and buying power of the segments be measured (e.g., number of customers)?
2. *Substantial*: Are the segments large and profitable enough to serve?
3. *Accessible*: Can the segments be effectively reached?
4. *Differentiable*: Do the segments respond differently to products and different marketing programs?
5. *Stable*: Are the segments sufficiently stable to justify the investment required to serve them (Blocker and Flint 2007)?

Of these five criteria, three are particularly relevant for breeding programs. First, a breeding program needs to know the size of different market segments in order to set breeding priorities. At a minimum, it needs to know *how many* farmers are involved in each market segment. Without this, resources may be invested in developing new products for a small minority of farmers. Second, it needs to know if these market segments differ in terms of the product profile that they require. Do all market segments have a unique set of trait preferences, or can one product profile meet the needs of more than one segment? Finally, what is the future of these market segments? Are they growing or declining? Which segments are growing fastest and would give the highest returns to new products? Are some market segments of producers responding to new market demands and are there opportunities for the breeding program to address these?

These five criteria for market segmentation are designed to maximize market advantage for private companies, which prioritize market segments according to their commercial potential. They may find no advantage in targeting the poor who are by definition involved in production for home consumption, are in thin markets, and play marginal roles in value chains. Instead, multinational seed companies target growers of hybrid seeds with access to agro-dealers and output markets. These five criteria therefore need to be adapted in order to target poor producers. The priority is to identify which market segment must be targeted in order to achieve the biggest reduction in poverty. Measurability, size, and stability are relevant criteria. But accessibility (distance to markets) and size (in terms of profitability) are less relevant if the objective is to target the poor. Other interventions may be needed to help meet these criteria and ensure that the breeding program is effective.

Gender should be included as a criterion for market attractiveness. If women producers are a disadvantaged group, then market segments of women producers or segments that include high numbers of women producers cannot be evaluated simply according to market criteria of size and buying power. Otherwise, the program may end up ignoring women, or marginalizing them further.

To be gender responsive, therefore, the breeding program must also take gender equity into account in deciding which groups to target.

The target segments are homogeneous in that they express a specific set of market demands. To be useful to breeders, these demands must be prioritized by the target segment (by ranking or scoring) so that breeders know the relative importance the market attaches to individual traits. Moreover, the reasons behind these preferences have to be known. For example, women growers may express a preference for leafy vegetables, because the leaves shade out weeds and save labor for weeding. Thus, a preference for leafy vegetables is explained by a labor constraint on women's time. Framing market demands as constraints and not just as trait preferences opens the door to other solutions apart from breeding. In the case of leafy vegetables, herbicides may offer a better solution for weed control than plant breeding.

Step 6: Identify which and how many segments to target

The decision on which segments to target depends on the objective and resources of the breeding program, the feasibility of the breeding required, and the importance of the segment. Some segments may be highly attractive but cannot be served because the program lacks the necessary competencies to serve them effectively. In contrast, a market segment with trait preferences that are difficult to breed for may be targeted because the potential impact is highly desirable (e.g., when a breeding program decides that a transformative objective is required to meet its goals). Thus, targeting a market segment involves strategic choices in which gender considerations should be one of the decision-making criteria (Kotler 2000). These choices are discussed in more depth in section 2.3.

2.2 Positioning

Positioning is "the act of designing the company's offering and image to occupy a distinctive place in the target market's mind" (op. cit.). There are two steps:

- *Step 7: Develop a positioning strategy.* This involves differentiating the product from rival products in the mind of the customer.
- *Step 8: Design appropriate marketing mix to communicate positioning.* Differentiating the product allows the company to decide on the marketing mix or how to persuade consumers to buy the product. The marketing mix—product, price, place, and promotion—drives the communication strategy for a specific product.

Increasingly, Steps 7 and 8 have become the responsibility of public sector breeding programs. For example, the uptake of biofortified crops, like orange-fleshed sweetpotato, and the emphasis on sorghum and millets as "smart foods" rely heavily on advertising and the media to create consumer demand. The growing emphasis on nutrition in public sector breeding programs will require greater attention to market positioning. Steps 7 and 8 (*market positioning*) are important for products with "hidden" traits that are new to producers, such as biofortified products or products with high levels of vitamin A. The growing emphasis on nutrition has made market positioning an important step in public sector breeding programs. (That said, positioning is not discussed further in this paper, whose main focus is on market segmentation and targeting.)

2.3 INCORPORATING GENDER INTO SOCIAL TARGETING

When incorporating gender into social targeting, breeding programs will decide whether to take a functional or transformative approach to gender, especially when formulating breeding objectives. In practice, these decisions arise repeatedly during segmentation, targeting, and positioning. To illustrate, the matrix in Table 3 depicts four options that can occur when deciding on a market segment to target, taking gender equality into account. Each option entails either a functional or transformative approach to incorporating gender into targeting.

Options 1 and 2 involve a functional approach to gender, where changes to gender equality are not the motivation for varietal improvement but considering gender can potentially improve adoption and impact. Option 1, where there is no change to existing breeding products, is a continuation of the status quo: when, for example, targeting analysis establishes that there are no economically important gender inequalities that can be addressed though varietal improvement. This involves a functional approach to gender equality which, after proper consideration, is found to provide no new opportunity for program impact.

By contrast, in Option 2 targeting analysis finds there is opportunity for new, more relevant products that can be tailored to needs shaped by existing gender relations. For example, if women cannot access timely irrigation and there is a call for new varieties with improved tolerance to early-season drought. Breeding objectives and positioning of a new product are shaped by understanding differences in gender roles, but there is no intent to change these roles (see Box 2). The decision to target women's needs increases the scope of adoption and impact for the breeding program. This is a functional approach to gender which involves adaptation to existing gender relations that continue to dictate unequal resources or different roles. Gender equality might be reduced once women have access to a new breeding product; yet this is not the primary objective of varietal improvement.

Options 3 and 4 illustrate the decision to breed for a target segment for which changes in gender equality are an objective. In Option 3 targeting analysis establishes that important gender inequalities cannot be addressed cost-effectively through breeding, as in Option 1. The difference is that breeding objectives in Option 3 are designed in conjunction with other transformative interventions (usually implemented by other actors). For example, varieties already under development are positioned in a different way, packaged with improved access to fertilizer, credit, or small machinery designed to transform use of a variety, previously beyond the scope of a target group of women in this market segment.

In contrast, Option 4 can involve a decision to undertake development of a new breeding product that is inherently transformative. For example, this may be a decision to breed explicitly to meet demand from women for specific traits (e.g., earlier maturation, new storage or processing qualities) that will enable these producers to shift to a new production frontier or enter new markets, deliberately inducing a positive change in existing gender inequalities. Additionally, varietal improvement may be designed to be explicitly contingent on the introduction of other innovations, such as new processing technology, intended to transform existing gender relations. Both Options 3 and 4 put the onus on the breeding program to innovate with the intention of effecting changes in gender equality.

Table 3. Examples of decisions that incorporate gender into targeting through a functional or transformative approach to gender equality

		GENDER EQUALITYIN THE MARKET SEGMENT	
		EXISTING GENDER RELATIONS	**NEW (MORE EQUAL) GENDER RELATIONS**
BREEDING PRODUCTS	**EXISTING PRODUCTS**	Option 1: Existing products for use under existing gender relations	Option 3: Existing products for use under changed (more equal) gender relations
	NEW PRODUCTS	Option 2: New products target improved relevance under existing gender relations	Option 4: New products that change (increase equality of) gender relations
		Functional approach	*Transformative approach*

Source: Authors.

Box 2. Precooked beans in Kenya: New products target improved relevance under existing gender relations

CIAT asked Lasting Solutions Inc., a consumer marketing company, to help identify the market for precooked beans in Kenya. The company first defined what they wanted to achieve, which was to improve nutrition among the poor in Kenya by increasing the consumption of beans rich in iron and protein. Next, they studied how different types of households bought, cooked, and ate beans. They discovered the importance of gender roles in bean consumption: women were always responsible for preparing and cooking beans, and urban shoppers were men but rural shoppers were women. Women in urban areas only become shoppers once they reached a higher standard of living.

Understanding these gender roles helped the marketing team to identify whom they needed to influence if they wanted to change consumer behavior to accept a new product. The final step was to understand consumer preferences: what distinctive features did shoppers look for when they bought beans? The team discovered that shoppers don't buy nutrition: they buy beans according to "looks," primarily color. The trick was to identify and promote beans with high protein and iron that had the right color. They also discovered that, because fuel was expensive, precooked beans could halve the cost of cooking beans. This increased the amount of money people could spend on beans and would help increase bean consumption, which would in turn helps to improve nutrition.

Acknowledgment: Thanks to Joab Ouma of Lasting Solutions Inc., for sharing his experience with us.

At what stage in the breeding program is targeting information needed?

Stage 1, *segmenting the market*, needs to be done very early in the breeding cycle: for setting the most basic breeding objectives in relation to the size and economic significance of one or more populations of end-users and/or subpopulations within them. Typically, breeders identify market segments based on biophysical factors (e.g., length of growing period) or on a widely prevalent production constraint (e.g., drought) within a particular agro-ecosystem. They then identify the "must-have" traits that will allow a new variety to adapt to this agro-ecosystem. But this ignores the implications of diversity among end-users, for some of whom a different constraint may be a higher priority. Programs do not usually characterize the market in terms of market segments of producers, except in very general terms ("resource-poor farmers," "subsistence farmers").

Stage 2, *targeting specific market segments*, should also happen early in the breeding program, when breeders build a business case for their products. At this stage, programs need to prioritize their client group of resource-poor farmers. They may draw on information about trait preferences provided by social scientists. In most public sector breeding programs, however, targeting is ad hoc and does not involve priority-setting according to the size of the market segment based on the number of resource-poor growers or potential size of the market. Targeting also occurs in the *development* stage of the breeding program, when breeders test their products in farmers' fields. Products are market-tested using, for example, participatory variety selection, whereby different groups of end-users (farmers, women, processors) provide feedback to the program.

Stage 3, *positioning the product*, in which advertisers design a marketing mix that will appeal to the target group, occurs when the breeding product is being commercialized. This is not usually part of a public sector breeding program.

Where does gender and social targeting fit in this process? Table 4 juxtaposes stage-gate[2] and the STP framework. This shows that Steps 1–6 of the framework relate to the *discovery* phase of the stage gate process. Step 1 (*define the market*) forms part of stage-gate 1 (*scoping*), whereas Steps 2–6 form part of stage gate 2 (*build the business case*). Hence, the STP framework mainly belongs to the design stage of the breeding program, before product development.

In practice, gender and social targeting may take place when products have already reached the development phase. Information on trait preferences is often collected from farmer participatory trials when new products are being tested. This information is used to confirm that the program has identified the priority traits correctly, and to test whether its products have these traits. It may only be at this stage that gender preferences are revealed. This means that gender and social targeting is ex-post, with researchers working backwards to fit products into market segments rather than vice versa.

Table 4. Information needs in a plant breeding program

Stage-gate and Description		**STP Framework**
Discovery Phase		
Scoping	Quick evaluation of technical merits and market prospects	Define the market
Build the business case	Technical, market, and business feasibility analyses	Select bases for segmentation Validate the segments Construct segment profiles Evaluate market attractiveness Identify market segments to target
Development Phase		
Development	The technology is developed and process plans for the development of the new products are mapped out	
Testing and validation	Plans/assumptions are tested at production/manufacturing, product design, market, and financial levels	
Commercialization Phase		
Launch	Full production	Develop positioning strategy Design appropriate marketing mix
Review		

Source: Authors.

[2] Stage-gate is a project management technique in which an initiative or project (e.g., new product development, software development, process improvement, business change) is divided into distinct stages or phases, separated by decision points.

3. WHAT INFORMATION IS/IS NOT AVAILABLE FROM LARGE DATASETS?

In this section we identify the large datasets that can be used for market segmentation, and discuss how breeding programs might use these datasets to identify market segments for a specific crop. We use the example of cassava in Nigeria to illustrate how these datasets can provide gender-disaggregated information that can help breeding programs incorporate gender into their customer profiles.

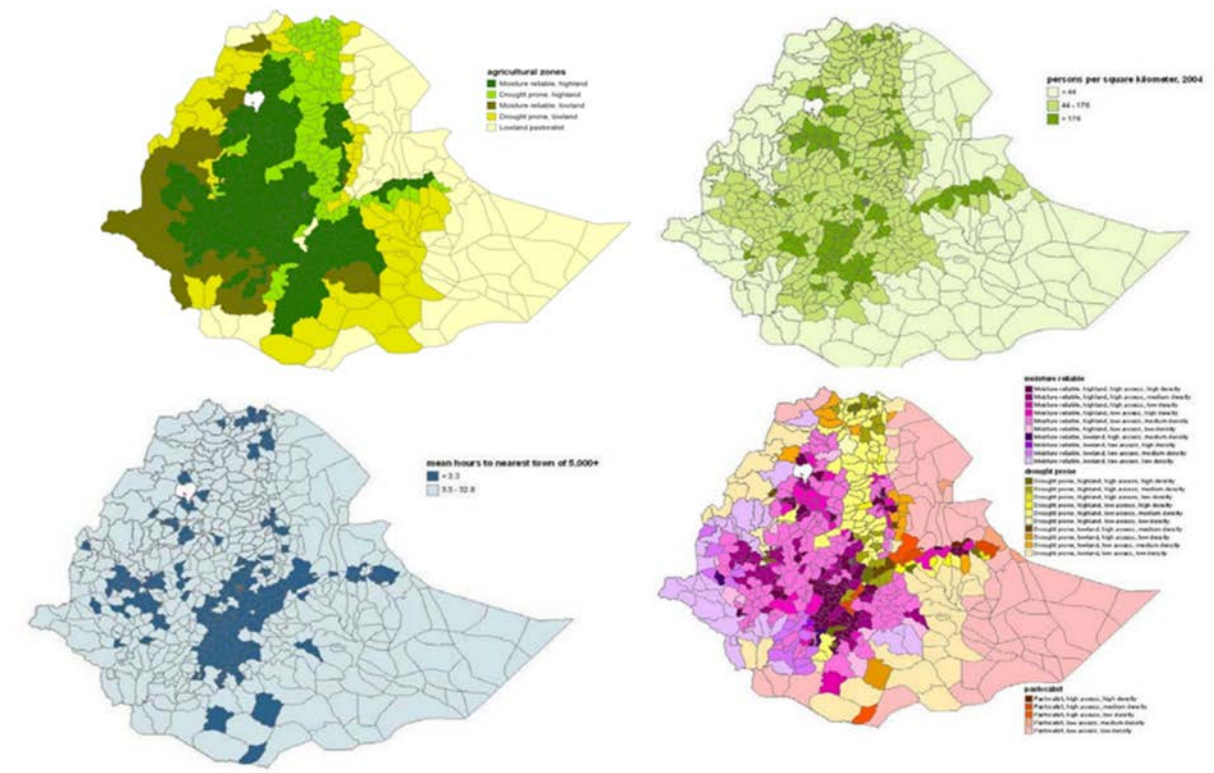
3.1 AN INVENTORY OF DATASETS

Choices for resource-poor farmers are strongly conditioned by the geography in which they live. The Ethiopian government, for example, frequently frames policy discussions by distinct geographical conditions of moisture (e.g., reliable, drought prone, and pastoral), otherwise known as the “Three Ethiopias” (Chamberlin et al. 2006). At the same time, other geographical factors such as poverty, population density, and access to resources are also critical in rural development. Impact studies for different kinds of investments in east African highland production systems, for example, have shown the importance of both biophysical and socioeconomic contexts (reviewed by Chamberlin, ibid.). Likewise, global analyses of farming systems illustrate both the necessity and possibility of mixed datasets (Dixon et al. 2001).

Multidisciplinary, big-data products increasingly underlie innovation targeting in agricultural development by honing-in on geographical hotspots.[3] For example, Chamberlin et al. (op. cit.) used spatial and country census datasets to segment the Three Ethiopias into more meaningful distinct geographical areas. In these areas agricultural conditions, constraints, and opportunities are relatively homogeneous and distinguishable from others, based on agro-ecology, access to markets, and population density (Figure 2). Similarly, the Association for Strengthening Agricultural Research in Eastern and Central Africa is upscaling agricultural technologies by segmenting their geographical domain using the same three variables and evaluating their impacts accordingly (Omamo et al. 2006; Gunaratna et al. 2010; Johnson and Flaherty 2010; Koo et al. 2016). Such examples, though relevant to the STP framework, are in a sense broad-based “first-order strategy” filters, whereas implementing specific technologies, such as variety-specific interventions, require more specifically defined spatial targeting frameworks (Chamberlin et al., op. cit.).

[3] Here, the term big-data refers to large-scale (multiple countries, region wide, or globally available), geotagged, and open-source datasets such as spatially explicit rasterized and vector data, remotely sensed satellite imagery, geospatial information systems, and nationally representative household survey data (i.e., microdata).

Figure 2. Row-wise from upper left: Agricultural potential zones, optimal thresholds for population density, selected market access indicators, and resulting segmentation for Ethiopia.



Source: Chamberlin et al. (2006).

At the same time, there has long been a data divide between the natural and social sciences. Although environmental data have successfully entered the "cosmopolitan age"—that is, data without borders, facilitated in part by technological advances in the data sciences, such as remotely sensed satellite imagery, geospatial information systems (GIS), and computer modeling—socioeconomic data are often framed by national boundaries (Otto et al. 2015). A shifting paradigm in data-gathering over the past 15 years, however, has steadily improved access to detailed socioeconomic datasets, thanks in part to a growing global alliance increasingly investing in microdata collection, cross-country standards, open data policies, and data-visualization platforms. As a result, subnational socioeconomic data products are increasingly available to the public, such as population and poverty grids[4], microdata derived from national household surveys[5], and rasterized sociodemographic indictors[6] (Azzarri et al. 2016) (see Appendix 1 afor examples of subnational datasets). These products are often overlooked in the economic literature, but they are well suited to the study of crop and human geography across scales.

[4] For example, WorldPop (http://www.worldpop.org.uk).
[5] For example, World Bank Integrated Surveys on Agriculture (LSMS-ISA) (http://go.worldbank.org/BCLXW38HY0) and the Demographic and Health Surveys Program (http://dhsprogram.com/data).
[6]For example, IFPRI's CELL5M geospatial database: a product of the HarvestChoice project (International Food Policy Research Institute and University of Minnesota); access latest CELL5M datasets from Dataverse (https://dataverse.harvard.edu/dataverse/harvestchoice).

Both mappable and microlevel data on income and poverty, population density, nutrition and health, and market access (to name a few), and when combined with biophysical datasets, are well suited for a STP breeding framework at scale (Appendix 1). Rasterized data are particularly useful for segmentation, allowing for easy aggregation of grid cell-level information across space. The CELL5M geospatial database[7] developed by the HarvestChoice[8] project at the International Food Policy Research Institute (IFPRI), for example, contains more than 750 harmonized data layers on agriculture, agro-ecology, demographics, and markets across sub-Saharan Africa (SSA) at 5 arc-minute (approx. 10 x 10 km^2) resolution (Table 5) (Koo et al. 2016). By overlaying spatially explicit, biophysical data layers with socioeconomic data, it is possible to investigate complex relationships between population and the environment across relevant geographical boundaries (e.g., watersheds, farming systems, or climatic zones). To illustrate such already ongoing analyses, Azzarri et al. (2016) presented a series of maps that integrate biophysical datasets with bottom-up data pooled from georeferenced household surveys, showing a spatial relationship between agro-ecological zones (AEZ) and early childhood wasting in SSA (Figures 3 and 4). The interoperability and capacity of mixed datasets, however, depend on many factors, including sampling methodology, data quality, and geographical coverage, as well as overlapping primary sampling units.

[7] The CELL5M geospatial database is available from Dataverse, https://dataverse.harvard.edu/dataverse/harvestchoice?q=cell5m

[8] All HarvestChoice data are available via Dataverse, https://dataverse.harvard.edu/dataverse.xhtml?alias=harvestchoice

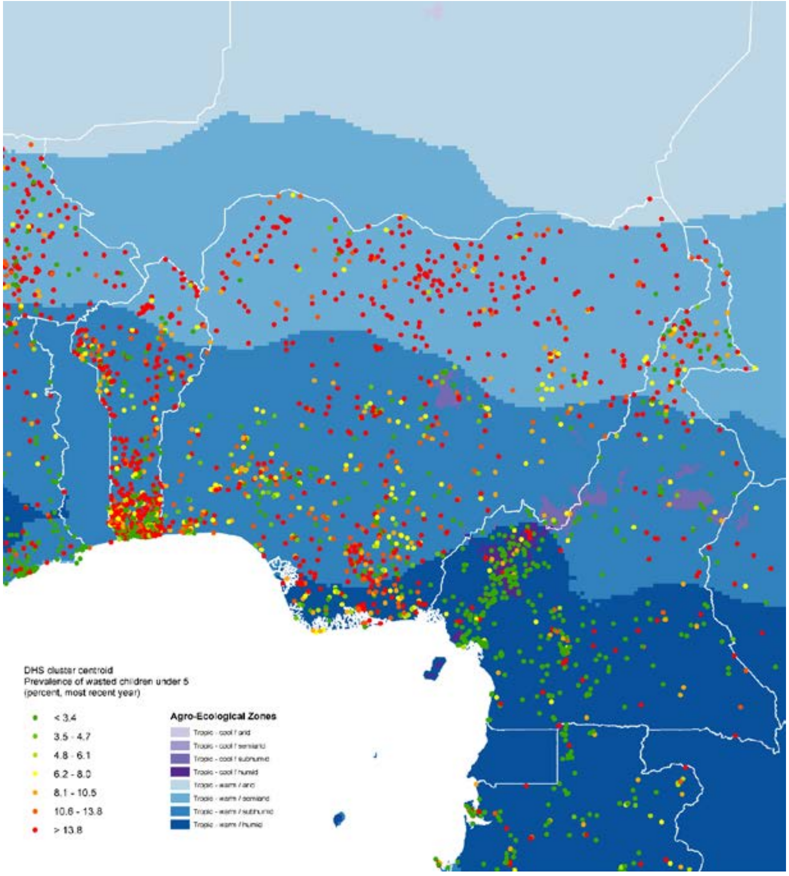
Table 5. CELL5M data layers by category, subcategory, and quantity thereof (as of April 2016)

Category	Subcategory (Number of data layers)
Agriculture	Harvested Area of Crops (134) Crop Production (134) Value of Crop Production (134) Crop Yield (134) Crop Yield Variability (2) Livestock (16)
Demographics	Health and Nutrition (90) Income and Poverty (36) Population (12)
Agro-ecology	AEZ (4) Climate (7) Elevation (1) Farming Systems (2) Land Cover and Land Use (21) Pests and Diseases (8) Soil Resources (19)
Markets	Marketshed (1) Portshed (1) Travel Time (11)

Note: HarvestChoice's CELL5M geospatial database covers agriculture, agro-ecology, demographics, and market access for SSA at 5 arc-minute resolution (10 km^2) available from Dataverse, https://dataverse.harvard.edu/dataverse/harvestchoice.

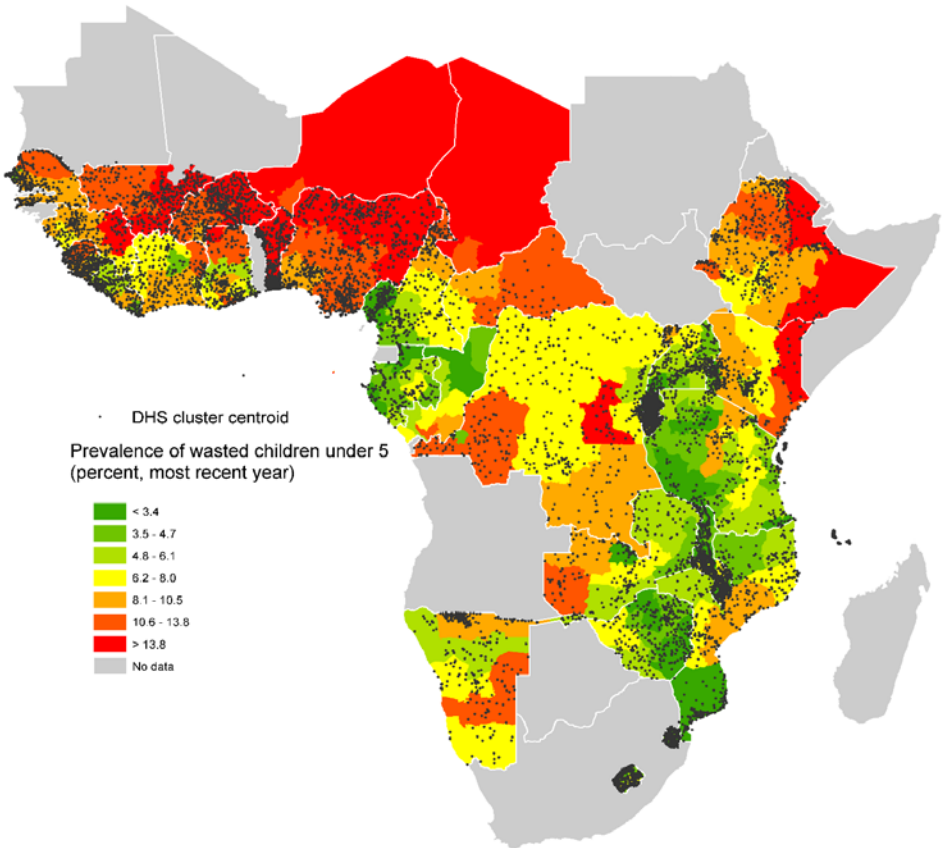
Source: Koo et al. (2016).

Figure 3. Centroids of Demographic and Health Surveys (DHS) clusters overlaid onto AEZ in Nigeria and surrounding countries. The color of each DHS cluster indicates the prevalence of wasted children under the age of 5 years. DHS data are nationally representative; AEZ data are from the CELL5M (see Appendix 1 and Table 5).



Source: Azzarri et al. (2016).

Figure 4. Centroids of DHS clusters overlaid on the estimated prevalence of wasted children under 5 years at the subnational level. Cluster GPS coordinates in Niger, Burkina Faso, Gabon, and Congo are not available. Gray areas indicate countries not covered by DHS data or with missing information on wasting.

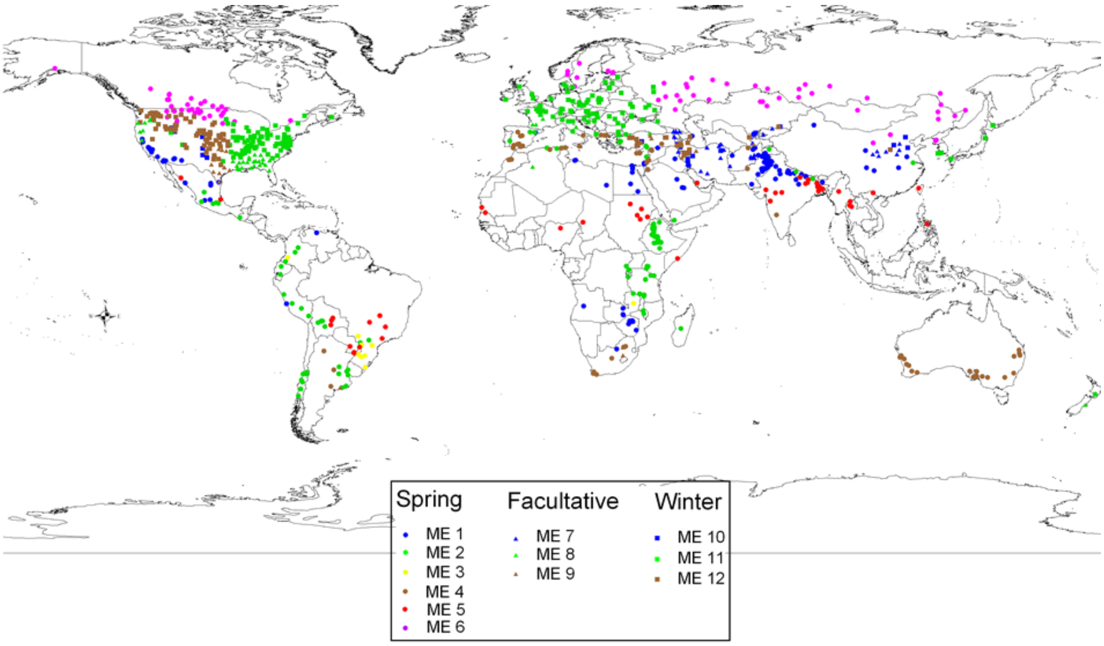


Source: Azzarri et al. 2016

In a plant breeding program, germplasm or variety targeting must operate within geographical factors imposed by the environment. The first order in market conglomeration is, therefore, generally determined by the crop. For example, to address the needs of diverse wheat growing areas, CIMMYT uses the concept of mega environments (ME) to target germplasm development (Rajaram et al. 1993; Hodson and White 2007).[9] A ME is defined as a broad, not necessarily contiguous area, occurring in more than one country and frequently transcontinental. In addition, a ME is defined by similar biotic and abiotic stresses, cropping system requirements, consumer preferences, and by a volume of production (Figure 5).

[9] http://wheatatlas.org/megaenvironments

Figure 5. Wheat ME



Source: Hodson and White (2007)

At the same time, breeding programs are normally framed by a specific area or geopolitical boundary. For example, IITA's cassava breeding for Nigeria is largely mandated at the national level, traditionally focusing on the development of widely adapted varieties. And although political constraints may pose a challenge by limiting options for varietal diversity, environmental parameters do not preclude the STP framework. Rather, such biophysical considerations are consistent with early steps (1–2) described in the segmentation process, representing one dimension of a larger targeting framework that includes socioeconomic and demographic data such as gender-sensitive and poverty variables.

3.2 Using large datasets to apply the STP framework

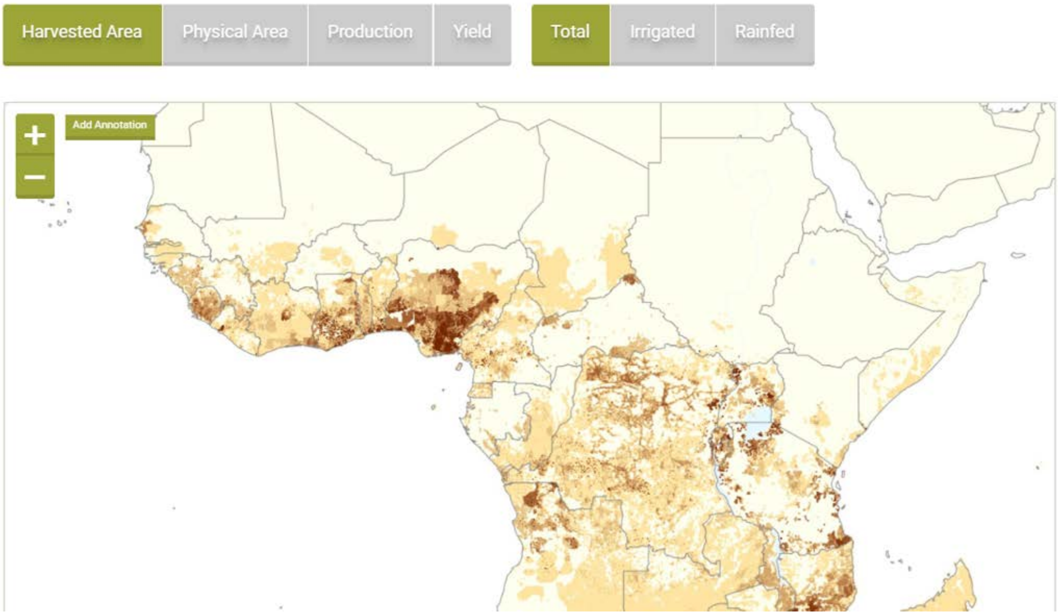
A market segment is "A group of producers with homogeneous trait preferences for a breeding product, taking into account the gender differentiation of preferences" (Table 1). The first step in identifying a market segment, therefore, is to characterize the population of potential customers, using a mix of biophysical and socioeconomic variables. The second step is to link this characterization of a group of producers to a set of homogeneous trait preferences. This second step is discussed in section 3.3.

3.2.1 Segmentation—Step 1: Define the market

In the context of a breeding program, "the market" is defined as the users of a breeding product within a relevant country or region. Depending on program objectives, value-added traits may target one or more players along the value chain, from seed disseminators and farmers to processors, traders, and consumers. At the farm level, various crop-specific, agro-ecological data are available for meaningful market segmentation within a crop's wider geographical boundary, such as soils, climate and crop suitability (current and future scenarios), volume of production, yield, and harvest area

(Appendix 1). Moreover, a combination of such factors can provide input to create more elaborate cropping information systems. For example, IFPRI's Spatial Production Allocation Model (SPAM)[10] uses a variety of inputs such as population density and crop-specific suitability information based on local landscape, climate, and soil conditions, producing plausible estimates of global crop distribution within disaggregated units. Measurements for 42 crops (90% of the global share of crops) are available in terms of four variables—area harvested, physical area, production, and yield—and two production systems, irrigated and rainfed (Figure 6) (You et al. 2014).

Figure 6. Cassava total harvest area from the SPAM, 2005 v. 2.0. Darker colors indicate higher values.



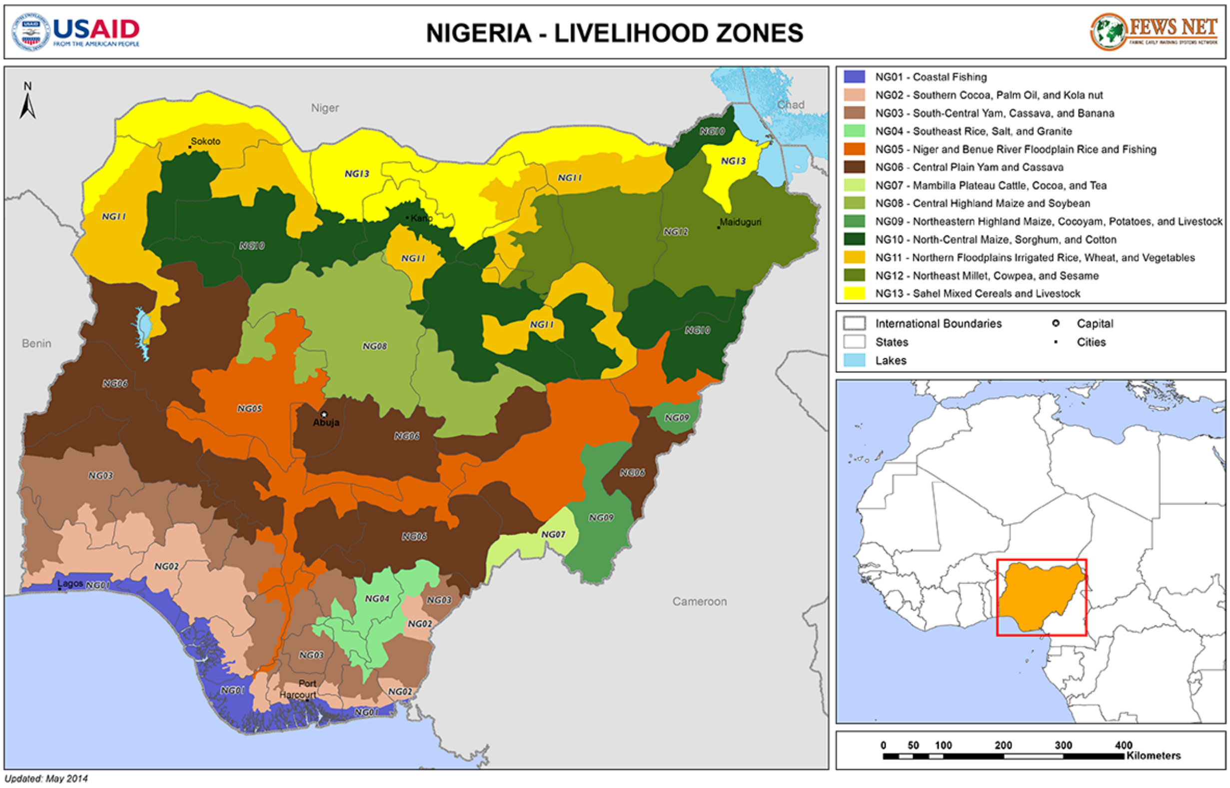
Source: http://mapspam.info/maps/. Updated datasets available from Dataverse. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DHXBJX

FEWS NET[11] data on livelihood zones may offer additional layers of geographical information relevant to the crop. Livelihood zones use data on agro-climatology, elevation, land cover, market accessibility, sources of food, and major economic activities to define geographic areas where people generally share similar options for obtaining food and income and similar access to markets (Figure 7). As such, livelihood zones offer a meaningful level of aggregation relevant for variety-specific targeting at the country level. Alone, SPAM and FEWS NET data may lack key geographical information. For example, SPAM may pinpoint areas of high volume production for a crop, whereas livelihood zones reflect the crop's relative economic importance for livelihoods and food security. By

[10] IFPRI and International Institute for Applied Systems Analysis (2016). "Global Spatially-Disaggregated Crop Production Statistics Data for 2005, Version 3.1." doi:10.7910/DVN/DHXBJX, Harvard Dataverse, v. 8.
[11] FEWS NET data available at http://www.fews.net/west-africa/nigeria/livelihood-zone-map/may-2014

overlaying these datasets, along with expert consultation and ground-truthing as well as statistical validation, we can begin to form more meaningful geographical segments at the farm level.

Figure 7. Livelihood zones in Nigeria

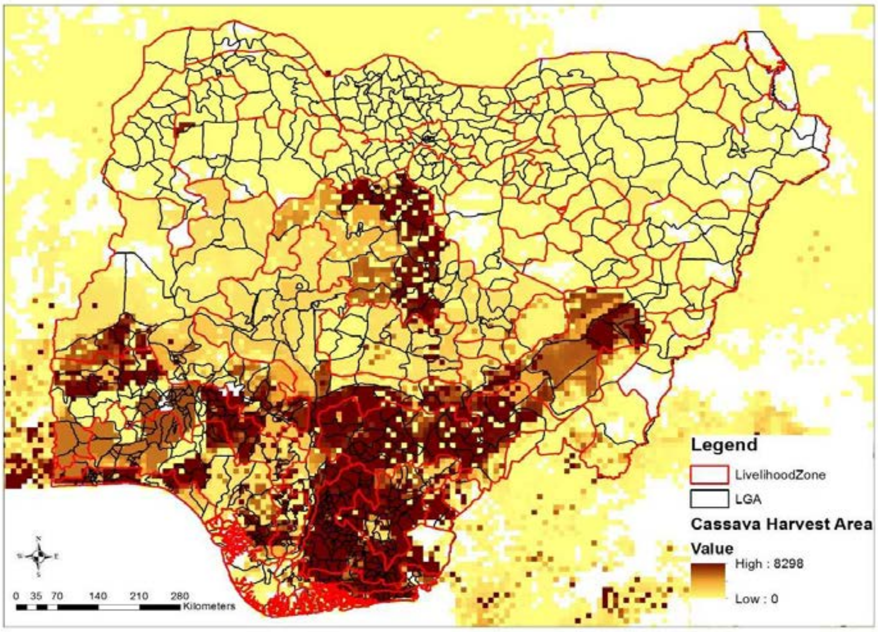


Source: FEWS NET, downloaded September 14, 2017, from http://www.fews.net/west-africa/nigeria/livelihood-zone-map/may-2014

For illustrative purposes, take the case of cassava germplasm or variety targeting in Nigeria, the world's largest producer of cassava. Cassava is a woody shrub extensively cultivated as an annual crop in tropical and subtropical regions for its edible and starchy tuberous root. It is a popular source of carbohydrates known for its low input requirements and drought tolerance. Cassava is also largely considered a "woman's crop" in SSA (Forsythe et al. 2016). To define and segment the market of growers within national boundaries, we first consider the geographical information regarding the crop by overlaying data onto cassava harvest area from SPAM and livelihood zones from FEWS NET, as discussed above (Figure 7).

According to the map (Figure 8), cassava cultivation is clearly most important in the southern half of Nigeria, the distribution of which is influenced by, for example, terrain, rainfall, and the Niger and Benue rivers, which converge and empty into the Niger Delta. As expected, cassava harvest area is relatively low in the most southerly area of Nigeria near the gulf, where tree crops and fishing are primary sources of food and income. Cassava production is also scarcer in the north, where dryland cropping and livestock are more suitable under low rainfall conditions. River proximity and flooding throughout the country favor rice production and fishing activities. Although cassava is a widely adapted crop cultivated and consumed throughout Nigeria, prime cassava producing areas are in the south-central and central plain regions according to both datasets. Accordingly, it is possible to begin segmenting the cassava market in Nigeria based on this agro-ecological assessment, identifying differential areas of cassava production within “simplified” livelihood zones (through redistricting or elimination of segments as informed by cassava statistics), and validated through statistical rigor and expert ground-truthing (Step 3). Note that while such analyses are relevant for farm-level targeting (growers), there is a dearth of information necessary for targeting processors and consumers, for example.

Figure 8. Cassava harvest area from the SPAM, 2005, v. 3.1 overlaid with Nigeria livelihood zones from FEWS NET



Source: Authors.

3.2.2 Segmentation—Step 2: Select bases for segmenting the market

As demonstrated above, the primary basis for segmentation is often geographic. Segmentation based on environmental parameters alone, however, are too coarse for breeding programs designed to target resource-poor farmers or consumers as well as others along the value chain. Demographic and socioeconomic data on rural populations, poverty, market access, as well as plot-specific microdata on end-uses and gender, for example, can inform the next level of market disaggregation, using openly available datasets previously mentioned (Appendix 1). For example, population density and market access are common bases for segmentation in SSA. Population density reflects available labor that may drive uptake of labor-intensive or land conservation practices, whereas market access is a sometimes unpredictable although useful variable to determine market opportunities and input use (Chamberlin et al. 2006). As previously discussed, Chamberlin and cohorts used mixed datasets to identify optimal thresholds for population density and selected market access indicators within homogenous AEZ, resulting in 25 segments appropriate for a broad-based policy framework in Ethiopia (Figure 2). Market access was represented by high and low mean woreda travel time (> and <3.3 hours) to nearest town of 5,000 or more inhabitants. Population density was divided into three classes of high (greater than 176 persons/km^2), medium (44–176 persons/km^2), and low (fewer than 44 persons/km^2). The authors adjusted variable parameters within fractionated segments to maximize their explanatory power by looking at the responsiveness of different outcomes and by testing the amount of variance of key rural livelihood indicators based on different thresholds. Using the cassava example illustrated above (Figure 8), we can then further segment geographical segments using similar socioeconomic variables and available datasets. Moreover, microdata derived from nationally representative household surveys on plot- and crop-specific gender variables, for example, can provide insights into gender dynamics within geographical segments of the population (Appendix 1).

Data derived from multidisciplinary surveys that cover a large geographical extent—for example, the Living Standards Measurement Study–Integrated Surveys on Agriculture (LSMS–ISA)[12], the Women's Empowerment in Agriculture Index (WEAI)[13], and the Demographic and Health Surveys (DHS) program[14]—are potentially rich sources of socioeconomic and gender-sensitive data, particularly in SSA (Appendix 1). These surveys are repeated over time and geocoded at the plot and/or household level.

DHS are nationally representative, population-based surveys with large sample sizes for more than 90 countries, offering valuable information on health and nutrition of women and children. Such information is critical for targeting poor consumers and improving nutrition, such as with biofortified crops. Indicators are presented in terms of national-level statistics and for population subgroups such as those defined by age, education, marital status, economic status, urban/rural residence, and region of the country.

DHS collects representative data in four areas: population, health, HIV, and nutrition.

The WEAI measures the empowerment, agency, and inclusion of women in the agriculture sector to identify ways to overcome those obstacles and constraints within the 19 focus countries of the U.S. Government's Feed the Future program. At the time this report was written, datasets from pilot questionnaires were openly available for three countries (Uganda, Bangladesh, and Guatemala). Subsequent country surveys are not yet openly available for online distribution, although presumably they can be recovered on request. It is important to note, however, unlike other household- and individual-level datasets mentioned here, WEAI data are not nationally representative but, rather, focus on the United States Agency for International Development's Zones of Influence within Feed the Future countries.

WEAI tracks women's agricultural engagement in five areas: production, resources, income, leadership, and time use.

LSMS–ISA surveys are perhaps the most comprehensive agricultural survey for specific crops. They include data on individual crop/plot managers and decisionmakers, and is thus highly compatible with the STP breeding framework. ISA surveys can be a tool for understanding gender dynamics in low-income settings, through the emphasis on collecting data on and from individuals in the household. ISA data address multiple topics, allowing for detailed analysis of the linkages between welfare, agriculture, and income diversification in SSA. ISA data are also disaggregated at the individual and farm-plot levels, enabling analysis of a wide variety of issues from a gendered perspective. Individual information is collected on who is responsible for the management and

---

[12] World Bank Integrated Surveys on Agriculture (LSMS-ISA) available at http://go.worldbank.org/BCLXW38HY0
[13] Women's Empowerment in Agricultural Index available at https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/19237
[14] DHS program available at https://dhsprogram.com/

decision-making of agricultural land. Furthermore, all ISA surveys have a panel component. The eight country datasets released to date (for Ethiopia, Niger, Nigeria, Malawi, Tanzania, Uganda, Burkina Faso, and Mali) cover at least half of the population of SSA, although questionnaires and the sampling framework vary between countries. For example, although all ISA surveys are representative at the national-level, survey panel data for Nigeria are nationally representative based on 5,000 households, which are also representative of the six geopolitical zones (at both urban and rural level) in Nigeria. In other words, statistical power may weaken at the level of state or local government area compared with larger aggregations.

The LSMS-ISA collects household and agricultural-related data in seven areas:

- Land inventory
- Crop production
- Crop management
- Labor and time allocation
- Inputs and improved varieties
- Resources
- Animal holdings.

As stated, ISA data are disaggregated at the household level, as well as at individual and plot/crop levels. This allows for crop-specific gender dynamic analyses potentially important to a crop breeder. Tuning into survey key words such as "who" and "ID code" provides a quick return of possible gender-specific information imbedded within LSMS–ISA surveys. As an example, consider the available variables in the 2010/11 Malawi Third Integrated Household Survey:

- **Plot details**. Plot-level information is recorded for each plot owned or cultivated by a farm household. Data (e.g., gender, age, education, and position in the family) are collected on and from individuals in the household who are at least partially (or jointly) responsible for the plot, via the household Roster ID code. Information is also collected on plot size and tenure, soil characteristics, crops planted and harvested, decision-making, time allocation, and household and hired labor at different stages of the season.
- **Coupon use**. Information is collected on each household member on coupon use, including who received a coupon, type of coupon (e.g., urea, maize seed, DAP), input purchasing with coupons, reasons for non-redemption, and the like.
- **Sales/Storage**. Information is collected on earning decisions within the household and identifies who/what the buyer/outlet is for crop sales.
- **Livestock**. Information is collected within households on livestock ownership (identifying individual household members or joint responsibility), management (feeding/care-taking), and earning decisions from livestock products.
- **Extension**. Information is collected on all household members in terms of agricultural advice/ information received through identified sources, including frequency and perceptions of quality.

Although there are no large-scale datasets that capture farmers' crop-trait preferences across populations, ISA data on crop-specific, individualized questionnaires can potentially capture key information on cassava growers and end-uses (e.g., Who manages the plots planted to cassava in Nigeria? Is cassava truly a woman's crop in SSA?) that is valuable to a cassava breeding program. Using geocoded household data on individual cassava-plot managers from the most recent Nigeria LSMS-ISA and overlaid with the map in Figure 9, we can pinpoint where in Nigeria women and men are growing cassava, and if there is spatial variation that can be explained by other geographical factors. Figure 9 shows a concentration of women cassava growers in the south-central area of Nigeria, in and around the Delta states. Cassava growers in the central and more northerly region of the country, as well as the southwest, are almost exclusively men. The spatial pattern of cassava growers follows a similar distribution of cassava harvested area from SPAM and livelihood data from FEWS NET. More analyses are needed, however, to segment Nigeria by other socioeconomic variables such as poverty and population. (The end result should be a manageable number of market segments with maximum explanatory power.)[15] The gender of decisionmakers during the postharvest phase of the cropping cycle was also identified from survey data (e.g., who decides what to do with the crop harvest, and who controls the earnings from any sales). Although many data values were missing, the distribution of men and women followed a similar pattern as in Figure 9 (not shown). As an illustration, we further overlaid extreme poverty data from WorldPop[16] on previous map layers and asked the question: Where are the poor women cassava farmers relative to men? From the map (Figure 10) we see that although much of Nigeria is poor, and both men and women cassava farmers are in areas of high and low poverty, female cassava plots are more concentrated in the poorer southeast region than in the southwest.

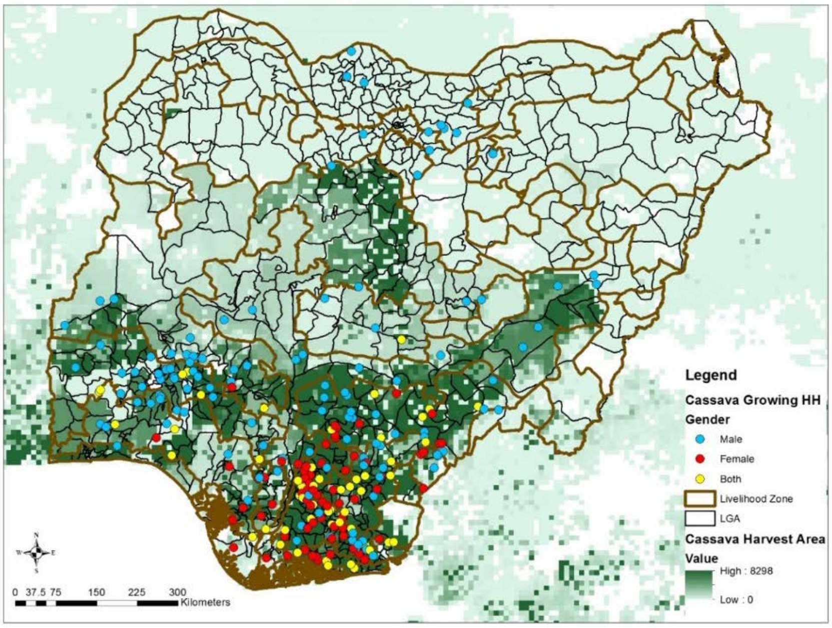


Figure 9. Gender of cassava plot managers using LSMA-ISA panel wave 3 data (2015–2016) for Nigeria and overlaid with cassava harvest area and livelihood zone data from SPAM and FEWS NET, respectively. Gender is determined by plot-manager ID information for cassava plots reported for each household and mapped using household geocoordinates.

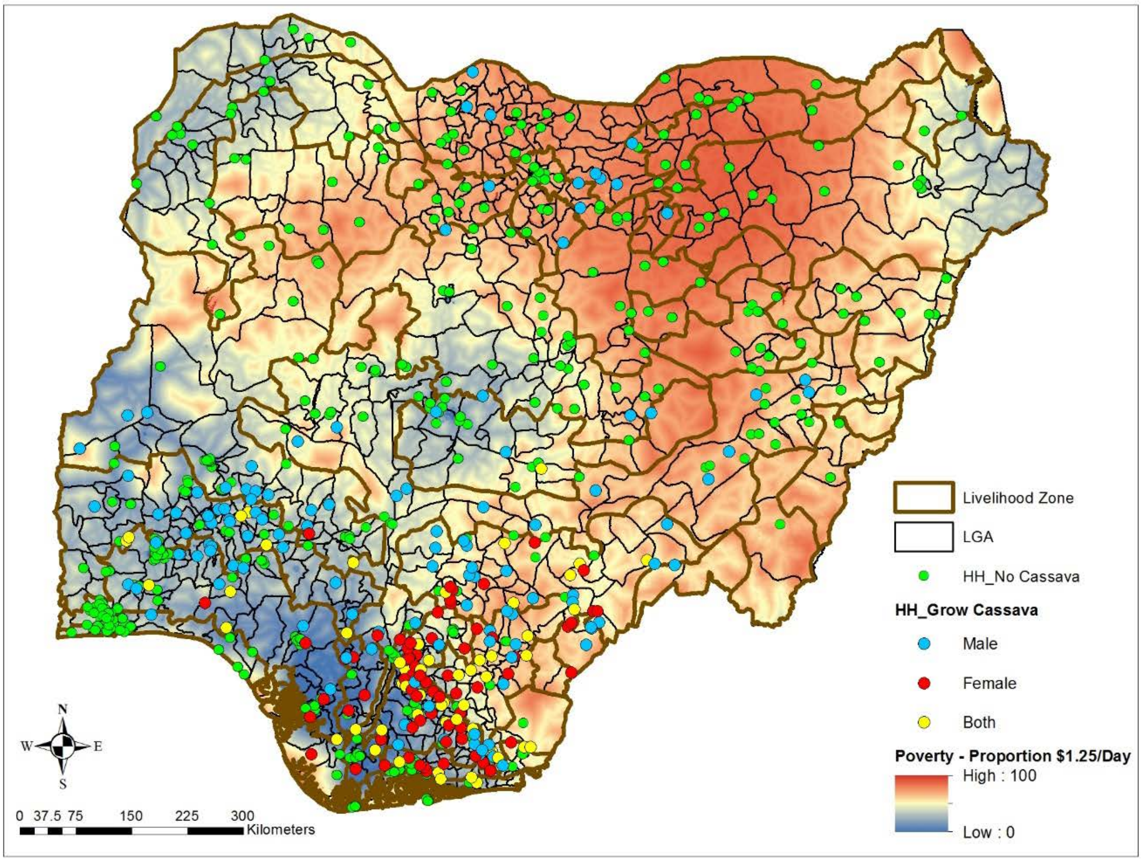
Source: Authors.

---

[15] Note that detailed plot information on labor and decision-making may add additional information regarding women's responsibility in household cassava production, both during post-planting and postharvest activities.

[16] WorldPop data available at http://www.worldpop.org.uk). Data are based on 2005 and reference outdated poverty lines. 2010 data may be available from PovcalNet and HarvestChoice's CELL5M geospatial database: http://iresearch.worldbank.org/PovcalNet/povOnDemand.aspx/ https://dataverse.harvard.edu/dataverse.xhtml?alias=harvestchoice

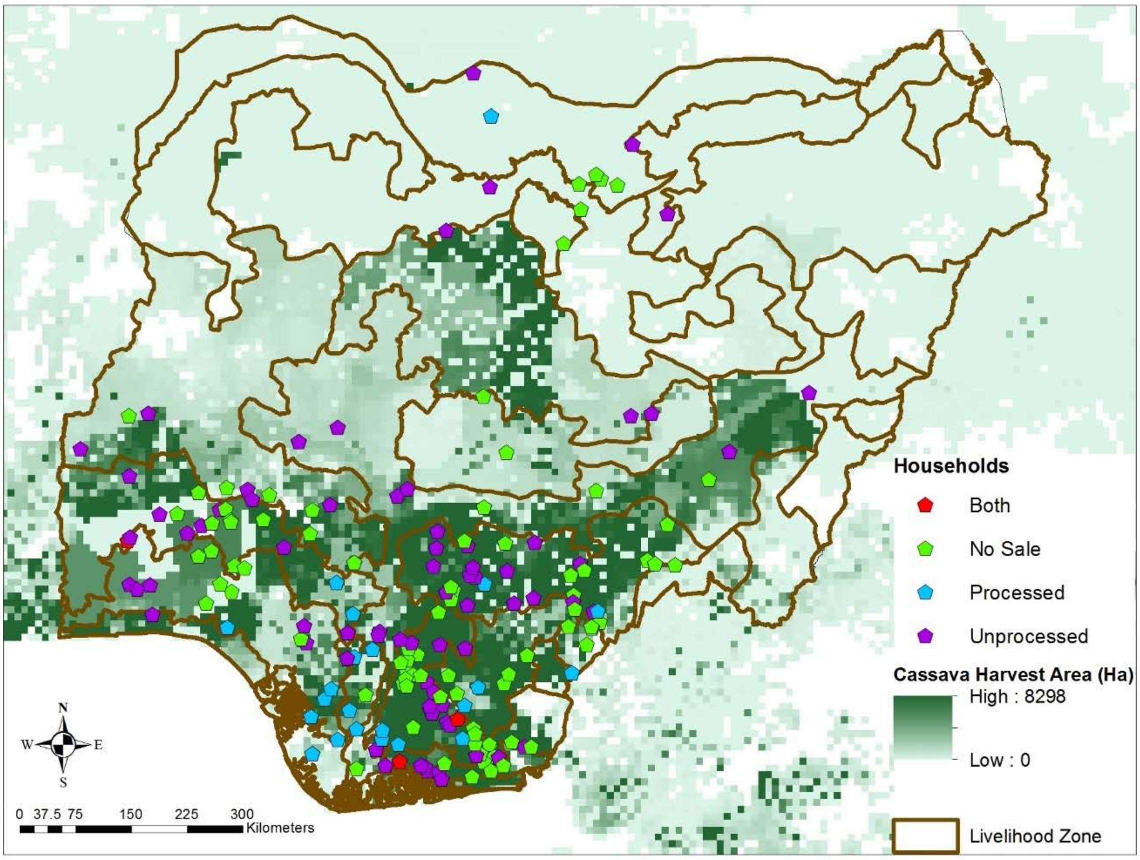
Figure 10. Cassava-growing households and gender of plot managers overlaid with extreme poverty.



Source: Authors.

Although the maps above (Figures 9 and 10) reveal gender differences and appear to follow a spatial pattern, there is virtually no information about cassava trait preferences; we can only rely on assumptions and proxies with these data (discussed further in section 3.3). Moreover, the data are more relevant early in the value chain at farm level, lacking detailed information once the crop leaves the farmgate. For example, women cassava processors are important in Nigeria, where various value-added cassava products are available to consumers. A clear understanding of gender roles is therefore needed before using gender as a basis for segmentation. By at least pinpointing the geography of men and women producers in relation to other geographical factors, however, we can begin to elaborate more on the differences between and within genders—both before and during the targeting stage of the STP. For example, Figure 11 shows the spatial distribution of household use of harvested cassava; that is, whether households kept the harvest on-farm or sold either processed or unprocessed (or both) cassava products. According to the map, Nigerian farm households generally do not engage in the sales of both unprocessed and unprocessed cassava; rather they focus on one or the other. At the same time, many households throughout Nigeria do not engage in the sale or trade of cassava harvests. Instead, they retain the harvest for various uses such as home consumption or livestock feed—variables that are embedded within the ISA data. Additional variables available from ISA survey data could perhaps provide a proxy of trait preferences, whether gender sensitive or not (e.g., farm and family size), if the plot manager is the primary decisionmaker regarding sales and use of the crop, the nature of household crop dependency (subsistence or cash), and the income and food expenditure of the household, to name a few.

Figure 11. Harvested cassava use among households. Households reported no sale (i.e., harvest stays on-farm) or sale of either processed or unprocessed cassava harvest.



Source: Authors.

Mappable data are necessary for segmenting a larger geographical area into homogenous groups with similar conditions, opportunities, and constraints. On the other hand, geocoded microdata collected by household surveys are potentially well suited for overlay with spatially explicit datasets, as well as providing data needed for descriptive summaries within populations and profiling segments during the targeting stage of the STP.

3.2.3 Targeting—Steps 4–6: Construct segment profiles, evaluate market attractiveness of segments, and identify which and how many segments to target

There a number of relevant household- and individual-level questions to ponder:

- What is the size of the total area of cassava per plot and/or per farm household?
- What are the characteristics of the plot?
- Who manages the crop on each plot, and what is their gender and position within the family?
- Who decides what crops to plant per plot and what to do with the harvest and earnings?
- Did the plot receive inputs or improved crop varieties?
- How much harvest did the household yield?

- What did the household do with the harvested crop?
- What is the socioeconomic status of the plot manager or household?

A breeding program's objectives will clearly depend on the populations intended for targeting, ultimately determining the bases for market segmentation and targeting as framed by the STP. It is important to note that most policy analysts recognize that one-size-fits-all strategies are inadequate for advancing development objectives at the national level. Yet at the same time, how much heterogeneity should be addressed at strategic planning levels? If blanket solutions are unreasonable, so too is the development of strategies for every household or community situation. Moreover, breeders must make long-term investments and decisions regarding a finite set or combination of genetic traits for a presumably diverse mix of farmers operating under different environments. A feasible STP breeding framework therefore cannot rely on too many—or too few—segmented markets for targeting. Rather, it should operate within manageable geographical parameters, considering the genetic and time constraints inherent in germplasm development, while capturing the enormous spatial heterogeneity of farmers and end-users into a simplified number of homogenous groups in ways that are both gender and poverty inclusive. In the end, careful consideration of the value chain is necessary, as growers are driven by personal preferences, the consumer audience, and market considerations, to name a few. And although big-data products are here to help, they cannot take the place of participatory research, such as participatory plant breeding and varietal selection, to truly capture gender-responsive trait preferences among targeted groups of resource-poor farmers and consumers and others along the value chain.

3.3 LINKING CHARACTERIZATION OF SEGMENTS TO HOMOGENEOUS TRAIT PREFERENCES

This section presents a proposed approach for combining characterization of market segments with the generation of data on trait preferences. The approach draws on discussion of how to approach sampling for targeting at the workshop on design principles for gender and breeding organized by the CGIAR Gender and Breeding Initiative (GBI 2017). This type of approach is especially useful when geographic and demographic data are adequate for identifying and validating market segments but behavioral data are lacking, especially data on trait preferences. Assume that we have used large-scale geographic and demographic datasets to identify different groups of producers who are reasonably homogeneous in terms of their geographical location and socioeconomic status.

These groups represent the potential customers for the breeding program. But they do not yet represent a market segment. To qualify as such, they must also be homogeneous in terms of their trait preferences. In other words, to describe producer groups as customers, they must want the same product. Thus, the second step in identifying market segments is to link these producer groups with a homogeneous set of preferences that specifies the product for the breeding program. This should be done using sampling to ensure that gender differences to be addressed by breeding are representative of social target groups at national and regional scales.

The approach proposed involves the following:

1. For a crop and a well-defined region (or country), create a set of breeding materials—varieties and land races—that displays traits (i) of known importance to breeders, (ii) of known importance to farmers, and (iii) deemed of potential future importance by breeders or industry or farmers.
2. Use large-scale datasets at country or regional scale to segment the national or regional population into homogeneous social groups, using geographic and demographic variables.
3. Draw a sample of respondents or focus groups from these macro-level segments for rapid appraisal interviews to obtain qualitative data about production systems, current and future demand, problems around the crop of interest, gender roles and responsibilities, and determinants of technology choice. Use of multistage cluster sampling may be appropriate.
4. Use the combined macro-level data and qualitative data to (i) to test the validity of previously defined population segments and (ii) refine the definition and characterization of population segments.
4. Use the sample of respondents from macro-level segments to select representative male and female users with whom to conduct participatory evaluation of the set of varieties defined in Step 1. This evaluation is a diagnostic exercise to establish the relative importance of different traits to different kinds of users (Ashby 1990). Collect sex-disaggregated socioeconomic data to characterize each respondent in the participatory evaluations.
4. Analyze this information across and within segments to discriminate distinct, homogeneous sets of ranked trait preferences and the socioeconomic characteristics of the users who express a given set of preferences.
6. Use this information to (i) assess the importance of segments and prioritize those the program will target and (ii) generate customer profiles for the selected target segments.
7. Ground-truth the customer profiles.
8. Incorporate information on sets of trait preferences and associated customer profile into product profiling.

4. HOW SHOULD WE DESIGN A TARGETING STUDY FOR A BREEDING PROGRAM?

In this section we discuss what research to provide breeding programs with information on targeting should look like. Specifically, we discuss the conceptual framework for such a study, the research questions, and the data and tools that this study could use.

4.1 DEFINE THE RESEARCH QUESTIONS FOR TARGETING JOINTLY

Targeting studies are best conducted when breeders and social scientists work as one team rather than in separate disciplinary groups. The starting point for an effective targeting study is that team members agree on the problem. Compare these three questions, which are illustrative and not intended to be programmatic or represent the full range of questions that might be asked by a breeding program:

- Question A: Is there a trait that can help to develop a new product that will lead to a breakthrough in adoption and result in impact on an industrial scale?
- Question B: Is there a trait that can help to develop a new product that will lead to a breakthrough in adoption by **resource-poor farmers (including women)** that will result in impact on an industrial scale **and that may also benefit poor consumers?**
- Question C: Is there a trait that can help to develop a new product that will lead to a breakthrough in reduction of inequality in the capture of profits in the value chain between rich and poor producers—in particular, benefiting poor women in marginal production areas—and on an industrial scale?

Question A focuses on *what* and *how without explicit attention to* who or the distribution of benefits. Questions B and C incorporate explicit attention to *who.* Question B addresses the *who*, using a functional approach to gender equality, by including women in adoption but not addressing the distribution of benefits among gender-differentiated target groups. In contrast, Question C incorporates a transformative objective by asking whether there is potential for positive change for a specific target group. When the question of *who for* is not made explicit, breeding programs may develop the wrong products. Unless the question of *what for* and *how* are clear, targeting may miss the mark. Thus, the research question must be agreed jointly by breeders and social scientists at the very beginning of a targeting study.

4.2 USE THE STP FRAMEWORK

Most targeting studies will cover only one or two of the eight steps in the STP framework. Using the STP framework as a point of reference can help researchers see where their targeting study fits within the bigger picture of developing a targeted breeding program. The framework also reinforces the point that targeting studies are not an end in themselves but a means to an end, which is to improve the performance of the plant breeding program. The common goal of targeting studies is to enhance the impact of the program.

The STP framework is a formal, academic model of marketing that may not reflect actual practice (Dibb 1998; Simkin and Dibb 1998; Hoek et al. 1996). One way to make this framework more user-friendly is to rewrite it as a checklist. A checklist transforms the eight steps in the STP into a series of questions that require a “Yes” if the breeding program is to successfully meet the needs of resource-poor farmers. The checklist can help to identify potential weak points in the design of the breeding program and the type of information needed to fill gaps in knowledge.

Figure 12 suggests a checklist of key questions that breeding programs need to ask when developing products for resource-poor farmers.

Figure 12. A targeting checklist for breeding programs.

Segmenting the Market

- What types of producers will decide to adopt and use the product(s) of breeding?
- Has the program identified market segments in terms of geography/demography and behavioral variables using sex-disaggregated data?
- Has the program analyzed different types of producers within and across these market segments?
- Are there evident, economically or culturally significant gender differences that can be used to classify types of farmers?
- Is there a need and rationale for a transformative vs. a functional approach to gender as a program objective?
- Are some market segments not technically feasible/too costly for the program?

Targeting Market Segments

- Do we know the size, profitability, and number of male and female farmers in each segment?
- Have we used the right indicators to identify male and female resource-poor farmers?
- Do men and women play different roles in production/processing/marketing?
- Which market segments contain the majority of male or female resource-poor farmers?
- What are the trait preferences of these male or female farmers?
- What products are needed to match these trait preferences?

Positioning the Product

- How will male and female resource-poor farmers become aware of the product?
- Can they find it easily if they want to try it?
- How much will they pay to use it, and are there gender differences in willingness to pay?

Source: Authors.

Another way to operationalize the STP framework is to transform it into a matrix. Figure 13 shows a hypothetical example. The matrix is three-dimensional, with axes representing smallholders, end-uses for the crop, and gender. Smallholders are classified as poor and non-poor, to distinguish the client group for the breeding program. This example also shows how the concept of a value chain can be integrated into the STP framework. The crop has three end-uses: home consumption and two value chains where the crop is processed and/or sold. Finally, the third axis is gender, which allows for different gender roles in production and sale.

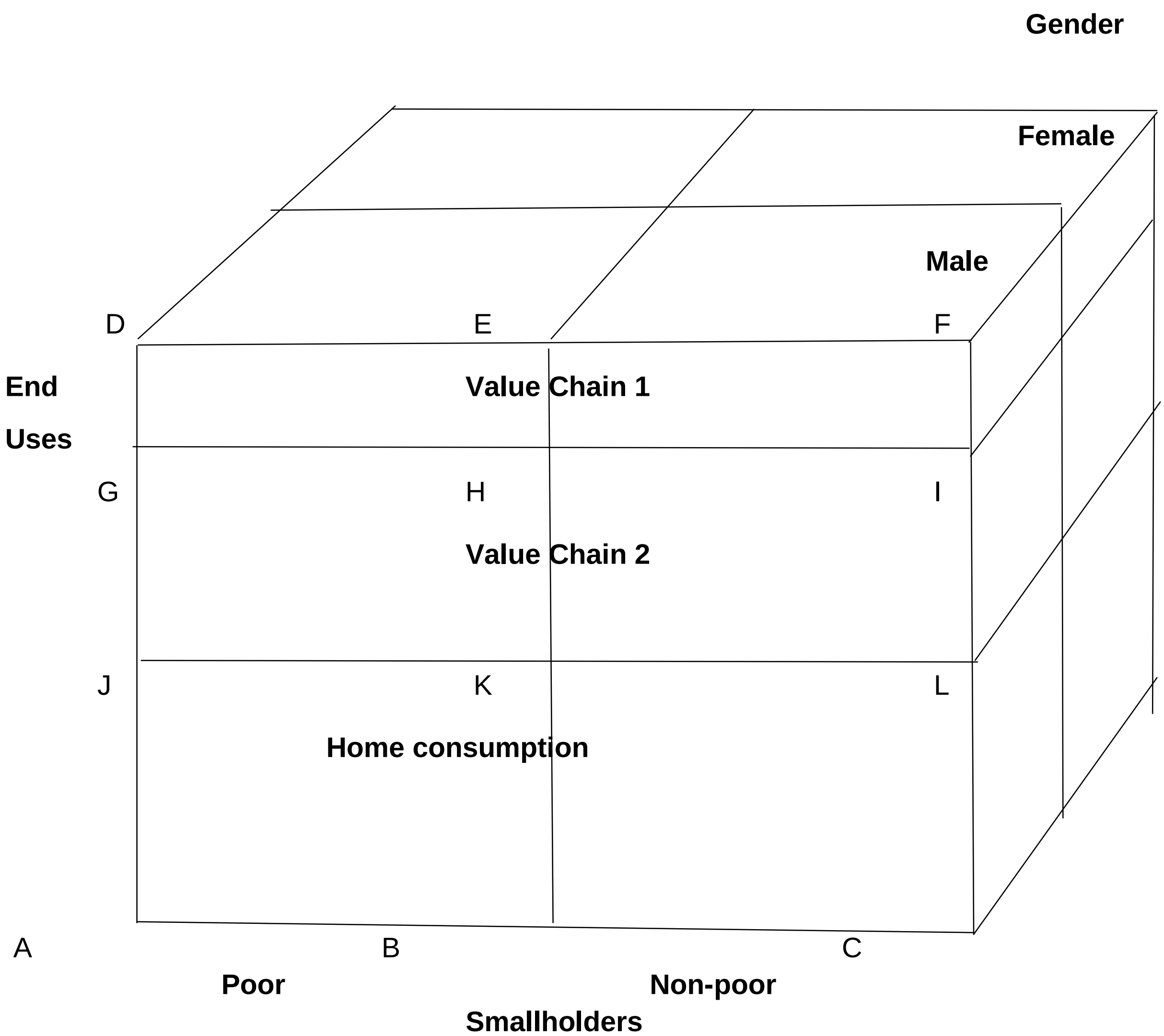


Figure 13. A hypothetical segmentation and targeting matrix.

Source: Adapted from Littler (1995).

The matrix can be used to show the steps from the STP framework to segment and target market segments:

- Step 1: Define the market, or the possible uses of the crop by all the growers. This is represented by the rectangle ACFD.
- Step 2: Select bases for segmentation (income, gender, end-use).
- Step 3: Validate the segments. The income and end-use variables give 6 market segments, whereas the addition of gender gives a total of 12 market segments.
- Step 5: Evaluate market attractiveness of segments. This is represented by the rectangle ABED, which are the segments that include the client group of resource-poor farmers.
- Step 6: Identify which and how many segments to be targeted. This requires choices between segments ABKJ (home consumption), JKHG (value chain 1), and GHED (value chain 2).

An example shows how this matrix might be used to segment the market for cassava in Nigeria. A recent study identified four value chains for cassava: (1) fresh roots by weight, (2) fresh roots by weight and variety, (3) processed cassava, and (4) fresh roots sold to the starch industry (Peters

2014). Value chain 1 involved a relatively small number of peri-urban producers, whereas value chain 4 involved larger farmers. The highest number of small producers was contained in value chain 3, where cassava was processed into *gari* for sale and into *akpu* for home consumption. Gender roles were important in this value chain because men produced the cassava while women processed the roots into *gari* and *akpu*. Thus, the main target segment for the cassava breeding program was value chain 3, and gender-responsive breeding required information from both men as growers and women as processors in order to determine the market demand for specific traits.

4.3 USE LARGE DATASETS

Section 3 discussed the large datasets that can give breeding programs information about resource-poor farmers at national level. Three key messages emerge:

- A minimum list of variables needed for targeting can be found in these large datasets. Information about trait preferences from small-scale studies can be linked to these datasets.
- Socioeconomic targeting precedes geographic targeting. Socioeconomic datasets are usually based on a sample of the wider population, so coverage varies across geographical or administrative regions. The first step, therefore, is to identify the growers and then link them to specific regions.
- Social targeting precedes gender targeting. The primary target group is resource-poor farmers; however, gender cuts across both the poor and non-poor. The first step, therefore, is to segment by income into poor and non-poor farmers and only then to segment resource-poor farmers by gender. Segmentation by gender assumes that there are major differences in trait preferences between men and women. If major differences do not exist, they can be treated as a single segment.

4.4 USE MIXED METHODS

The integration of quantitative and qualitative methods (Q-squared) is useful for gender and social targeting because, though qualitative methods give better insights into the reasons for trait preferences, quantitative methods are needed to segment the market into homogeneous groups at the national level and to develop customer profiles. A wide range of tools is available to improve targeting in breeding programs (Orr and Ashby 2016). Table 6 summarizes the information that is required and the relevant tools for each stage in the STP framework.

Table 6. Data and tools for targeting

STP Framework	Data Needs	Data Sources	Tools	Examples
Step 1. Define the market	Crop utilization Consumption patterns Income elasticity of demand Future demand	Agricultural statistics Industrial statistics Household expenditure surveys	Subsector analysis Value chain analysis IMPACT model (IFPRI)	Sorghum and millets in ESA (Gierend and Orr 2015; Orr et al. 2016a, 2017) Agri ProFocus (nd) Gender in value chains
	Crop harvested area Breeding environments Potential area	HarvestChoice	ArcGIS mapping of agro-ecosystems Crop suitability maps	Beans: Wortmann et al. (1998); Farrow et al. (2017) Malawi: Benson et al. (2016)
Step 2. Select bases for segmentation	Geographic (agro-ecosystem, distance from market) Demographics (income, gender, etc.) Behavioral (benefits, adoption status)	LSMS Baseline surveys (crop specific) “Immersion” studies	Cross-tabulation Correlation	Beans in ESA (Ouma 2016) Gender roles: Feldstein and Jiggins (1994); Andersson et al. (2016); Orr et al. (2016b)
Step 3. Validate the segments	Socioeconomic and gender indicators	LSMS Baseline surveys (crop specific)	Factor analysis Cluster analysis Discriminant analysis	Target group studies (Bidogeza et al. 2009)
Step 4. Construct segment profiles	Socioeconomic indicators not used as bases for segmentation	Same as Step 3	Cross-tabulation	Target group studies (Bidogeza et al. 2009)
Step 5. Evaluate attractiveness of market segments	Size of market segment Number of resource-poor farmers and women in each segment Growth and profitability of segments	Same as Step 3	Cross-tabulation	Cassava in Nigeria (Wossen et al. 2017)
Step 6. Identify which and how many segments should be targeted	Trait preferences for market segments Technical feasibility Research costs and benefits	Trait preference studies Adoption studies Expert opinion	Ranking and scoring; Conjoint analysis; Ex-ante cost-benefit analysis; Internal rate of return	Beans in ESA: Abeyasekera et al. (2002) Cassava: Bentley et al. (2017) Cowpea: Kristjanson et al. (2005)

Note: ESA = East and southern Africa.
Source: Authors.

Step 1: Defining the market

Several tools can provide quantitative data about the demand side of the breeding program (Lynam and Janssen 1992). These include subsector or value chain analyses, and national expenditure surveys, which provide information about household consumption by income, region, and for rural–urban areas. Foresight analysis using the IMPACT model (which includes crops and livestock) can project future demand based on income and population growth. The studies for sorghum and millets in ESA by ICRISAT show the type of information needed to define the market and quantify utilization (Gierend and Orr 2015; Orr et al. 2016a, 2017).

On the supply side, the market for a plant breeding program is also defined spatially using biophysical parameters. “Target population environments” (TPEs) cluster climate, soils, and management practices into unique environment groups, and trait packages are then developed for each environment (IRRI 2017). GIS tools can be used to develop TPEs, a crop atlas, and crop suitability maps to quantify the market for the plant breeding program (Benson et al. 2016; Wortmann et al. 1998).

Step 2: Select bases for segmentation

Income: Numbers of resource-poor farmers and poor consumers can be estimated from the LSMS, which provides information at the national level on own-production and consumption for each food crop. Accurate data on household income is hard to collect, so crop-specific baseline surveys (e.g., the Cassava Monitoring Survey) may not break down growers into poor and non-poor. Remember that resource-poor farmers are not a homogeneous group: to ensure food security, small-scale farmers in Nigeria plant three to four cassava varieties that mature at different times of year, whereas medium-scale farmers need fewer varieties.

Gender: Gender may not be a useful basis for segmentation since women are not a homogeneous group. In Nigeria, women’s roles in crop production, processing, and sales differ by religion (Muslim women cannot sell cassava in markets) and by income (women on larger farms are more involved in cassava processing than in production and may not have their own fields). Differences in trait preferences between men and women usually reflect differences in gender roles (see Box 3). Several tools exist that can provide information on gender roles (Feldstein and Jiggins 1994), including tools at the farm level (Orr et al. 2016a) and the level of the value chain (Andersson et al. 2016).

Box 3. Gender and trait preferences

Men and women play different roles in production, processing, and sales. Both tend to give higher priority to traits that are important to them as food-providers, or livestock-keepers, or processors.

In Nigeria, women give a higher priority than men to the ease of peeling cassava, whereas men give higher priority to the size of cassava roots. This is because women do the peeling, while men sell raw cassava. Women also prefer varieties that mature at different times because it gives them greater flexibility in the sale of processed cassava and allows a smooth flow of income throughout the year. Similarly, because women process cassava for sale, they prefer quality traits like softness and texture that command higher prices. This suggests that gendered trait preferences are complementary and reflect the gender division of labor.

A clear understanding of gender roles is therefore needed before using gender as a basis for segmentation. Once these roles are known, researchers can better understand why men and women may have different trait preferences, can identify which traits are important for different end-uses, and can then develop trait packages for each end-use. For example, if the market for cassava is defined and shows that *gari* is a major end-use for cassava, the breeding program can prioritize developing varieties that have these marketable traits. If a study of gender roles shows that (as in southwest Nigeria) men grow and women process cassava, then women’s trait preferences should receive attention when prioritizing the traits required for processing—not because they are women but because their gender role as processors gives them superior knowledge of the required traits.

Acknowledgments: Thanks to Peter Kulakow, Hale Tufan, and Bela Teeken for sharing their experience with cassava.

Geographic: Trait preferences vary in space not only because of differences in production environments (genotype x environment interactions) but also because consumer taste preferences may differ by region. In Nigeria for example, farmers in the southeast and northwest grow the same varieties of cassava but process them differently. Yet the sampling of LSMS surveys makes it difficult to align them with agro-ecosystems. Consequently, it is easier to use socioeconomic variables as the bases for segmentation, and then see whether there are spatial patterns and if they correlate with biophysical variables (e.g., market access).

After bases have been selected, statistical tools are used to split the sample into homogeneous segments. These tools are algorithms that form clusters based on the statistical distance between a set of selected indicators. The literature on target groups provides many examples (see Bidogeza et al. 2009 for references).

Step 3: Validate the market segments

Although statistical tools will always produce clusters of growers that are homogeneous in terms of the cluster variables, the clusters may not be useful for marketing purposes. They must be validated in terms of their size, accessibility, and stability, and whether they represent a distinct set of preferences.

At this stage researchers should test the meaning of key variables used to define the segments. For example, female-headed household (FHH) is often used as a proxy for gender; however, this indicator ignores heterogeneity among FHHs. Female heads who are single, separated, divorced, or widowed can have very different resources and capabilities. Incomes can also vary greatly since de facto FHHs (where the man is absent for 50% of the time) may have access to incomes from remittances (Chant 2004). Given this heterogeneity, what is the FHH variable actually measuring?

Trait preferences are best identified using qualitative methods such as focus groups, where preferences are scored and ranked (Abeyasekera et al. 2002). Focus groups can be designed to give a representative sample of growers (Bentley et al. 2017).

Step 4: Construct segment profiles

Once the clusters have been identified, the normal practice is to “profile” the clusters, or use variables that were not used in the original clustering to provide additional descriptive information about the cluster. This gives a clearer picture of the differences between the clusters and the distinctiveness of the segments. Again, the literature on target groups shows how this is done (see Bidogeza et al. 2009 for an example and further references).

Step 5: Evaluating market attractiveness

Whereas Step 3 focuses on the validity of the indicators themselves, Step 5 focuses on the resulting market segments and their potential value as a market for new products. This is measured by a range of criteria (size, accessibility, actionability, etc.). Another important criterion is stability, because it can determine whether the market segment is worth the investment. Segments and trait preferences can change over time. For example, poundability was once an important trait when maize was pounded by hand, but is no longer important since the introduction of hammer mills.

Step 6: Identify which segment(s) to target

A plant breeding program for resource-poor farmers must identify market segment(s) with two objectives in mind, market attractiveness and equity. This is the rationale for gender and social targeting. Only when these twin criteria are satisfied should the program consider technical feasibility. The program must decide which product lines give the highest payoff for its selected segment(s). These product lines are identified using the standard tools such as cost-benefit analysis and the internal rate of return.

5. CONCLUSIONS

Gender and social targeting can help breeding programs intended to benefit resource-poor farmers. The challenge facing these programs is to ensure that they have information about the trait preferences of resource-poor farmers, including women, and that this information is representative at the national level. This paper provides plant breeders with ideas, sources of information, and tools that social scientists can use to address this challenge.

Many of these ideas come from consumer marketing and are used by private sector breeding programs. By contrast, public sector breeding programs have development goals and are not profit driven. At the same time, however, they have to serve their customers; this requires effective marketing. Reconciling development goals with a marketing approach is a difficult balancing act. But since these programs have one foot in the marketplace, they can benefit by adapting ideas and tools from consumer marketing.

The main messages from this paper are the following:

- The STP framework provides breeding programs with a systematic approach to defining the market, identifying different market segments, and selecting which segments to target. This is the marketing approach used by commercial plant breeding programs. The framework also identifies the research questions, information, and tools needed at each stage of research process.
- Many large datasets that provide information about rural households at the national level are now available. Combined with datasets that provide biophysical data, they can provide breeding programs with information about growers, including their income, their numbers, their location, and how much they consume and sell. This set of variables can be used to categorize growers into homogeneous market segments.
- Some simple design guidelines can improve the relevance of targeting studies for breeding programs. These include joint design of the study by plant breeders and social scientists, using the STP framework to link the study to the objectives of the breeding program, providing information at scale, and the use of mixed methods to generate this information.

REFERENCES

Abeyasekera, S., Ritchie, J., Lawson-McDowall, J. 2002. Combining ranks and scores to determine farmer's preferences for bean varieties in southern Malawi. *Experimental Agriculture* 38:97–109.

Agri Pro Focus (nd). Gender in Value Chains. Practical Toolkit to integrate a gender perspective in agricultural value chain development. https://goo.gl/UHoLti

Andersson, K., Bergman Lodin, J., Chiwona-Karltun, L. 2016. Gender dynamics in cassava leaves value chains: The case of Tanzania. *Journal of Gender, Agriculture and Food Security* 1(2):84–109.

Aseete, P., Katungi, E., Bonabana, J., Ugen, M., Birachi, E. 2016. Ex-Ante Analysis of the Demand for new value added pulse products: A case of Precooked Beans in Uganda. Joint Pan-African Grain Legume and World Cowpea Conference 2016, AVANI Victoria Falls Resort and Conference Center, 28 February–4 March 2016, Livingstone, Zambia.

Ashby, J. 1990. Evaluating Technology with Farmers: a handbook. Investigación Participativa en Agricultura/ Participatory Research in Agriculture (IPRA), CIAT, Cali, Colombia. 95 pp.

Azzarri, C., Bacou, M., Cox, C.M., Guo, Z., Koo, J. 2016. Subnational socioeconomic dataset availability. *Nature Climate Change* 6:115–116.

Benson, T., Mabiso, A., Nankhuni, F. 2016. Detailed Crop Suitability Maps and Agricultural Zonation Scheme for Malawi. Spatial information for agricultural planning purposes. Feed the Future Innovation Lab for Food Security Policy. Research Paper 17. Washington, DC: International Food Policy Research Institute.

Bentley, J., A. Olanrewaju, T. Madu, O. Olaosebikan, T. Abdoulaye, T. Wossen, V. Manyong, P. Kulakow, B. Ayedun, M. Ojide, G. Girma, I. Rabbi, G. Asumugha, M. Tokula. 2017. Cassava farmers' preferences for varieties and seed dissemination system in Nigeria: Gender and regional perspectives. IITA Monograph, IITA, Ibadan, Nigeria.

Bidogeza, J.C., Berentsen, P.B.M., De Graaff, J., Oude Lansink, A.J.G.M. 2009. A typology of farm households for the Umutara Province in Rwanda. *Food Security* 1:321–335.

Blocker, C.P., Flint, D.J. 2007. Customer segments as moving targets: Integrating customer value dynamism into segment instability logic *Industrial Marketing Management* 36: 810–822.

CGIAR SRF. 2015. *CGIAR Strategy and Results Framework 2016-2030*. Montpellier, France.

Chamberlin, J., Pender, J.L., Yu, B. 2006. Development domains for Ethiopia: capturing the geographical context of smallholder development options. *IFPRI Discussion Paper: DSDG No. 43 and EPTD No. 159*.

Chant, S. 2004. Dangerous Equations? How Female-headed Households became the poorest of the Poor: Causes, Consequences and Cautions. *IDS Bulletin* 35(4):19–26.

Chege, C.K., Lungaho, M., Ndungu, J., Karanja, D., Birachi, E., Jager, M. 2016. Constraints of Bean Consumption by the Base of Pyramid (BoP) Consumers in Urban and Peri-urban Nairobi, Kenya. Joint Pan-African Grain Legume and World Cowpea Conference 2016, AVANI Victoria Falls Resort and Conference Center, 28 February–4 March 2016, Livingstone, Zambia. https://goo.gl/FWtD7k

Dibb, S. 1998. Market segmentation: strategies for success. Marketing Intelligence & Planning, Vol. 16 Issue: 7, pp. 394-406.

Dixon, J.A., Gibbon, D.P., Gulliver, A. 2001. *Farming systems and poverty: improving farmers' livelihoods in a changing world*. Rome: Food and Agriculture Association of the United Nations; Washington, DC: World Bank.

Farrow, A., Muthoni, R., Ojara, M. 2017. Updating the Atlas of Common Bean (*Phaseolus vulgaris*) Production in Africa: a tool for targeting and evaluating legume research and development. Joint Pan-African Grain Legume and World Cowpea Conference 2016, AVANI Victoria Falls Resort and Conference Center, 28 February–4 March 2016, Livingstone, Zambia.

Feldstein, H.S., Jiggins, J. 1994. *Tools for the Field. Methodologies Handbook for gender analysis in agriculture.* London: Intermediate Technology Publications.

Forsythe, L., Posthumus, H., Martin, A. 2016. A crop of one's own? Women's experiences of cassava commercialization in Nigeria and Malawi. *Journal of Gender, Agriculture and Food Security* 1:110–128.

[GBI] Gender and Breeding Initiative. 2017. Workshop Report.

Gierend, A., Orr, A. 2015. Consumer demand for sorghum and millets in eastern and southern Africa: Priorities for the CGIAR Research Programme for Dryland Cereals. Socioeconomics. Discussion Paper Series 35. Patancheru, Telengana, India: ICRISAT.Patancheru, Telengana, India: ICRISAT. http://oar.icrisat.org/9013/

Gunaratna, N.S., De Groote, H., Nestel, P., Pixley, K.V., McCabe, G.P. 2010. A meta-analysis of community-based studies on quality protein maize. *Food Policy* 35:202–210.

Haley, R.I. 1968. Benefit Segmentation: A Decision-oriented research tool. *Journal of Marketing* 32 (July):30–35.

Haley, R.I. 1984. Benefit Segmentation—20 Years Later*. Journal of Consumer Marketing* 1(2):5–13.

Hodson, D.P., White, J.W. 2007. Use of spatial analyses for global characterization of wheat-based production systems. *Journal of Agricultural Science*, 145, 115–125.

Hoek, J., Gendall, P., Esslemont, D. 1996. Market segmentation: a search for the Holy Grail? *Journal of Marketing Practice: Applied Marketing Science* 2(1):25–34.

[IRRI] International Rice Research Institute. 2017. The Target Population of EnVironments. https://goo.gl/qqtVbW

Johnson, M., Flaherty, K. 2010. *Strategic analysis and knowledge support systems for agriculture and rural development in Africa: Translating evidence into action* (Vol. 6). International Food Policy Research Institute, Washington, D.C.

Koo, J., Cox, C.M., Bacou, M., Azzarri, C., Guo, Z., Wood-Sichra, U., Gong, Q., You, L. 2016. CELL5M: A geospatial database of agricultural indicators for Africa South of the Sahara. F1000Research 5:2490.

Kotler, P. 2000. *Marketing Management*. The Millennium Edition. Englewood Cliffs, NJ: Prentice-Hall.

Kristjanson, P., Okike, I., Tarawali, S., Singh, B.B., Manyong, V.M. 2005. Farmers' perceptions of benefits and factors affecting the adoption of improved dual purpose cowpea in the dry savannas of Nigeria. *Agricultural Economics* 32:195–210.

Lipton, M. 2007. Plant breeding and poverty; can the Can transgenic seeds replicate the 'Green Revolution' as a source of gains for the poor? *Journal of Development Studies* 43(1):31–62.

Lipton, M., Longhurst, R. 1989. *New Seeds and Poor People*. London: Unwin Hyman.

Littler, D. 1995. Market Segmentation. In *Marketing Theory and Practice*, 3rd ed., M.J. Baker, ed. London: Macmillan, pp. 90–103.

Lynam, J.K., Janssen, W. 1992. Commodity Research Programs from the Demand Side. *Agricultural Systems* 39:231–252.

Omamo, S. W., Diao, X., Wood, S., Chamberlin, J., You, L., Benin, S., Wood-Sichra, U., Tatwangire, A. 2006. Strategic priorities for agricultural development in Eastern and Central Africa. Washington, D.C.: (Vol. 150) International Food Policy Research Institute (IFPRI).

Orr, A., Ashby, J. 2016. Gender-Responsive Plant Breeding: Applying Economic Tools and Methods. Paper presented at the Gender, Breeding and Genomics Workshop, October 18-21, 2016 Nairobi, Kenya.

Orr, A., Mwema, C., Gierend, A., Nedumaran, S. 2016a. Sorghum and Millets in Eastern and Southern Africa. Facts, Trends and Outlook. Working Paper Series No. 62. ICRISAT Research Program, Markets, Institutions and Policies. Patancheru, Telangana, India: International Crops Research Institute for the Semi-Arid Tropics. https://goo.gl/nvG3Hu

Orr, A., Homann-Kee Tui, S., Tsusaka, T.W., Msere, H., Dube, T., Senda, T. 2016b. Are there "women's crops"? A new tool for gender and agriculture. *Development in Practice* 26(8):984–997.

Orr, A., Gierend, A., Choudhary, D. 2017. Value Chains for Sorghum and Millets in Eastern and Southern Africa: Priorities for the CGIAR research program for Dryland Cereals. Socioeconomics Discussion Paper Series No. 42. Nairobi: ICRISAT. http://oar.icrisat.org/10007/

Otto, I.M., Biewald, A., Coumou, D., et al. 2015. Socioeconomic data for global environmental change research. *Nature Climate Change* 5:503–506.

Ouma, J. 2016. Linking Research to Product Development: A Case for New Beans Products in Uganda and Kenya. Joint Pan-African Grain Legume and World Cowpea Conference 2016, AVANI Victoria Falls Resort and Conference Center, 28 February–4 March 2016, Livingstone, Zambia. https://goo.gl/mXQiAR

Rajaram, S., VanGinkel, M., Fischer, R.A. 1993. CIMMYT's wheat breeding mega-environments (ME). In: *8th International Wheat Genetics Symposium*.

Simkin, L. and S. Dibb. 1998. Prioritising target markets", Marketing Intelligence & Planning, Vol. 16 Issue: 7, pp.407-417.7

Smith, W.R. 1956. Product Differentiation and Market Segmentation as Alternative Marketing Strategies. *Journal of Marketing* 21(1):3–8.

Sumberg, J., Reece, D. 2004. Agricultural Research through a 'New Product' Development Lens. *Experimental Agriculture* 40:295–314.

Sumberg, J., Heirman, J., Raboanarielina, C., Kaboré, A. 2013. From agricultural research to 'product development.' What role for user feedback and feedback loops? *Outlook on Agriculture* 42(4):233-242.

Tonks, D.G. 2009. Validity and design of market segments. *Journal of Marketing Management* 25(3-4): 341–356.

Wedel, M., Kamakura, W.A. 1998. *Market Segmentation: Conceptual and Methodological Foundations*. London: Kluwer Academic Publishers.

Weinstein, A. 2006. A strategic framework for defining and segmenting markets. *Journal of Strategic Marketing* 14(2):115–127.

Wortmann, C.S., Kirkby, R.A., Eledu, C.A., Allen, D.J. 1998. *Atlas of Common Bean (Phaseolus vulgaris L.) Production in Africa.* Cali, Colombia: Centro Internacional de Agricultura Tropical.

Wossen, T., Tessema, G., Abdoulaye, T., Rabbi, I., Olanrewaju, A., Alene, A., Feleke, S., Kulakow, P., Asumugha, G., Adebayo, A., V. Manyong. 2017. The cassava monitoring survey in Nigeria. Final report. IITA, Ibadan, Nigeria.

You, L., Wood, S., Wood-Sichra, U., Wu, W. 2014. Generating global crop distribution maps: From census to grid. *Agricultural Systems* 127:53–60.

APPENDIX 1. EXAMPLES OF AVAILABLE DATASETS FOR THE STP FRAMEWORK FOR BREEDERS

Stage of STP Framework	Variable	Dataset Name/Source	Description/Comments
Segmentation			
Step 1: Define the market(s) by crop and/or end- use			
	Agro-ecology and crop statistics:		
	Country/region	GADM database of Global Administrative Areas: http://www.gadm.org/.	Countries and lower level subdivisions Various file formats available for GIS
	Agro-ecosystem	Global Agro-ecological Zones (GAEZ), FAO/IIASA: http://www.iiasa.ac.at/web/home/research/researchPrograms/water/GAEZ_v4.html	Comprehensive information on inventory of land and water resources, evaluation of biophysical limitations, and production potentials of land
		HarvestChoice/IFPRI[1]: https://dataverse.harvard.edu/dataverse/harvestchoice	AEZ for SSA based on the methodology developed by FAO and IIASA. Dataset includes three classification schemes: 5, 8, and 16 classes, referred to as AEZ5, AEZ8, and AEZ16, respectively.
	Crop suitability	GAEZ, FAO/IIASA: http://www.iiasa.ac.at/web/home/research/researchPrograms/water/GAEZ_v4.html	A standardized framework for the characterization of climate, soil, and terrain conditions for analyzing synergies and trade-offs of alternative uses of agro-resources (land, water, technology) for food and energy production
		GLUES geoportal http://geoportal-glues.ufz.de/stories/globalsuitability.html	Similar to above, more recently developed global database at 30 arc-second
	Crop area	Spatial Production Allocation Model (SPAM)	Plausible estimates of crop distribution within disaggregated units at available for 42 crops, measured in terms of four variables: area harvested, physical area, production and yield, and two production systems: irrigated and rainfed
	Farming systems	GAEZ, FAO/IIASA: http://www.iiasa.ac.at/web/home/research/researchPrograms/water/GAEZ_v4.html; https://dataverse.harvard.edu/dataverse/harvestchoice	Spatial farming systems data in SSA per FAO's methodology; HarvestChoice gridded data available at 5 arc-minute resolution (10 km^2) for SSA
	Livelihood Zones	FEWS NET/ http://www.fews.net/fews-data/335	Geographic areas of a country where people generally share similar options for obtaining food and income and similar access to markets
	Soils	HC27 Generic Soil Profile Database: https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/20299	Seven soil profiles generated based on texture, rooting depth (proxy of water availability), and organic carbon content (proxy of fertility) and based on a meta-analysis of WISE 1.1 soil profiles in SSA
Step 2: Segment the market	Possible variables listed above and below		
	Demographic and markets		
	Poverty and income	HarvestChoice/IFPRI: https://dataverse.harvard.edu/dataverse/harvestchoice	Subnational poverty headcount ratios derived from nationally representative household surveys and population census information

Stage of STP Framework	Variable	Dataset Name/Source	Description/Comments
			poverty calculations are based on the comparison between the household per capita consumption expenditure (a synthetic indicator expressing the money-metric welfare utility level) and the $1.90 and $3.10/day poverty lines expressed in international equivalent purchasing power parity dollars in 2011
	Population	HarvestChoice/IFPRI: https://dataverse.harvard.edu/dataverse/harvestchoice	Spatial layers on population variables for SSA
		WorldPop: http://www.worldpop.org.uk	Archive of spatial demographic datasets for Central and South America, Africa, and Asia to support development, disaster response and health applications
		Gridded population of the World (GPW): http://sedac.ciesin.columbia.edu/data/collection/gpw-v4	Models the distribution of human population (counts and densities) on a continuous global surface
		UN World Population Prospect: https://esa.un.org/unpd/wpp/	A mix of historic and projected regional and national total populations from 1950 to 2050
	Health and nutrition	HarvestChoice/IFPRI: https://dataverse.harvard.edu/dataverse/harvestchoice	Spatial layers on nutrition, health, and dietary outcomes based on elaborations from the DHS.
		DHS: http://www.dhsprogram.com/	Designed to collect data on marriage, fertility, family planning, reproductive health, child health, and HIV/AIDS. Women of reproductive age (15–49) are the focus of the survey and women eligible for an individual interview are identified through the households selected in the sample. Consequently, all DHS surveys utilize a minimum of two questionnaires: household questionnaire and women's questionnaire.
	Market access	HarvestChoice/IFPRI: https://dataverse.harvard.edu/dataverse/harvestchoice	Spatial data on market sheds and distance to market for different population sizes and travel distances in SSA
	Micro-level crop- and gender-specific (survey) data		
	Area of crop planted by the household	The Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA):[2] http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTLSMS/0,,contentMDK:23512006~pagePK:64168445~piPK:64168309~theSitePK:3358997,00.html	Nationally representative community, household, and individualized surveys covering eight partner countries in SSA
	Share of crop sold/consumed by the household		Information is collected on earning decisions within the household and identifies who/what the buyer/outlet is for crop sales
	Gender of crop manager at plot level		Plot-level information on individuals in the household who are at least partially (or jointly) responsible for the plot, including gender, age, and position in the family, via the HH Roster ID code. Plot-level Information is also collected on plot tenure, plot characteristics, and household and hired labor at different stages of the season.
	Gender of decision maker at plot and crop level		Plot-level information on individuals in the household who are at least partially (or jointly) responsible for decision-making from planting through harvest and sales—including gender, age, and position in the family—via the HH Roster ID code

Stage of STP Framework	Variable	Dataset Name/Source	Description/Comments
	Access to extension services		Information is collected on all household members in terms of agricultural advice/ information received through identified sources, including frequency and perceptions of quality
	Health and nutrition	DHS: http://www.dhsprogram.com/	Designed to collect data on marriage, fertility, family planning, reproductive health, child health, and HIV/AIDS. Women of reproductive age (15–49) are the focus of the survey and women eligible for an individual interview are identified through the households selected in the sample. Consequently, all DHS surveys utilize a minimum of two questionnaires—a household questionnaire and a women's questionnaire
	Food Insecurity	Food Insecurity Experience Scale (FIES): http://inddex.nutrition.tufts.edu/guiding-framework/indicator/food-insecurity-experience-scale-fies	Experience-based food insecurity scales at household and individualized levels constructed from short questionnaires that capture households' behavioral and psychological manifestations of insecure food access; based on output from the model, households are classified along a spectrum of food insecurity, ranging from mild to severe. To date, FAO has released 2014 nationally representative FIES data for nearly 150 countries.
Step 6: Identify segments to target	Possible variables listed above and below		
	The following examples were extracted from LSMS-ISA surveys:		
	Coupon use (credits/loans)		Information is collected on each household member on coupon use, including who received a coupon, type of coupon (e.g., urea, maize seed, DAP), input purchasing with coupons, reasons for non-redemption, etc.
	Livestock		Information is collected within households on livestock (identifying individual household members or joint responsibility) on ownership, management (feeding/ taking care of), and earning decisions from livestock products
	Other household characteristics		Family size, farm and plot size and characteristics, wealth and income, adoption of improved crop varieties, consumption and food expenditures.

[1] HarvestChoice's CELL5M geospatial database covers agriculture, agroecology, demographics, and market access for SSA at 5 arc-minute resolution (10 km^2) available from Dataverse, https://dataverse.harvard.edu/dataverse/harvestchoice.

[2] Microdata availability focuses on the Living Standards Measurement Study–Integrated Surveys on Agriculture (LSMS-ISA) project, nationally-representative community, household, and individualized surveys covering eight partner countries in SSA (Burkina Faso, Niger, Ethiopia, Nigeria, Malawi, Tanzania, Mali, and Uganda). The project's primary objective is to foster innovation and efficiency in statistical research on the links between agriculture and poverty reduction in the SSA region. The LSMS-ISA supports multiple rounds of nationally representative panel survey with a multitopic approach designed to improve the understanding of the links between agriculture, socioeconomic status, and non-farm income activities.

Source: Authors.

The CGIAR Gender and Breeding Initiative brings together plant and animal breeders and social scientists to develop a strategy for gender-responsive breeding with supporting methods, tools and practices. The Initiative includes experts from across CGIAR centers and Research Programs, is coordinated by the CGIAR Research Program on Roots, Tubers and Bananas and the International Potato Center, and is supported by CGIAR Funders.